



HAL
open science

Gammacoronavirus aviaires : dynamique évolutive des populations virales au cours d'infections expérimentales chez les poulets et les dindes

Alexandre Flageul

► **To cite this version:**

Alexandre Flageul. Gammacoronavirus aviaires : dynamique évolutive des populations virales au cours d'infections expérimentales chez les poulets et les dindes. Sciences agricoles. Université de Rennes, 2020. Français. NNT : 2020REN1B045 . tel-03359760

HAL Id: tel-03359760

<https://theses.hal.science/tel-03359760>

Submitted on 30 Sep 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THESE DE DOCTORAT DE L'UNIVERSITE DE RENNES 1

ECOLE DOCTORALE N° 600

Ecole doctorale Ecologie, Géosciences, Agronomie et Alimentation

Spécialité : Microbiologie Virologie Parasitologie

Par

Alexandre Flageul

***Gammacoronavirus* aviaires : dynamique évolutive des populations virales au cours d'infections expérimentales chez les poulets et les dindes.**

Thèse présentée et soutenue à Ploufragan, le 17 décembre 2020

Unité de recherche : **ANSES Laboratoire de Ploufragan-Plouzané-Niort**
Unité Virologie Immunologie Parasitologie Aviaires et Cunicoles

Rapporteurs avant soutenance :

Astrid Vabret	Professeur, CHU de Caen
Mariette Ducatez	Directrice de Recherche, INRAE

Composition du Jury :

Présidente du jury :	Nathalie Dejuqc-Rainsford	Directrice de Recherche, IRSET, Rennes, France
Examineurs :	Nathalie Dejuqc-Rainsford	Directrice de Recherche, IRSET, Rennes, France
	Astrid Vabret	Professeur, CHU de Caen, Caen, France
	Mariette Ducatez	Directrice de Recherche, INRAE, Toulouse, France
	Sjaak De Wit	Professeur, GD Animal Health, Deventer, The Netherlands
Directrice de thèse :	Béatrice Grasland	Directrice des Recherches ANSES, Ploufragan
Co-directeur de thèse :	Paul Brown	Chargé de Projet de Recherche, ANSES, Ploufragan

Prélude

Ce manuscrit de thèse est le fruit de trois années de travail financées pour moitié par l'Agence Nationale de Sécurité Sanitaire, alimentation, environnement et travail (ANSES), pour moitié par la Région Bretagne, et portées par l'Université de Rennes 1. Sans ces acteurs, ce travail n'aurait pas pu être réalisé. Merci à eux.



Ces trois années se sont déroulées au laboratoire ANSES de Ploufragan-Plouzané-Niort, sur le site de Ploufragan, dans l'unité de Virologie Immunologie Parasitologie Aviaires et Cunicoles (VIPAC), après une courte période dans l'unité de Génétique Virale et Biosécurité.



Ce projet de thèse a été encadré par le Docteur Béatrice Grasland, en qualité de Directrice de thèse, et le Docteur Paul Brown, en qualité d'encadrant de thèse.

Remerciements

J'ai passé trois excellentes années. Avec quelques moments un peu stressants, il ne faut pas le nier, mais trois excellentes années. J'ai rencontré des personnes formidables, certaines marrantes, d'autres moins, mais toutes étaient très agréables. Ce fut trois années enrichissantes, qui sont déjà finies.

Alors merci !

Merci aux personnes qui ont accepté d'évaluer le travail accompli au cours de cette période de ma vie, les docteurs Nathalie Dejuqc-Rainsford, Mariette Ducatez, Astrid Vabret et Sjaak de Wit. Merci pour le temps que vous nous avez accordé, à moi, mais aussi à mes directeurs de thèse. J'espère que la présentation de ce travail sera aussi enrichissante pour vous qu'elle l'aura été pour moi.

Merci aux personnes qui ont composé mon comité de suivi de thèse, les docteurs Dominique Mahé, Aymeric Hans et Meriadeg Ar Gouilh, pour leurs temps, leurs avis et leurs conseils.

Merci aux personnes de GVB, même si je suis venu vous voir à plusieurs reprises, ce fût au final une courte période de travail à vos côtés.

Merci, merci, merci à l'équipe Bioinfo (Fabrice, Edouard et Pierrick) pour leur travail sur le pipeline et sa mise à disposition au public ! Merci !

Merci à Yannick, pour avoir été mon tout premier contact à l'ANSES de Ploufragan, et pour m'avoir accueilli une première puis une deuxième fois en GVB.

Merci à tous les membres parasito de VIPAC. Nous ne nous sommes que très peu côtoyé et c'est bien dommage.

Merci à tous les membres viro de VIPAC. J'ai passé trois excellentes années en VIPAC et cela c'est grâce à vous. Sans exception, vous avez tous été très accueillant. Vous m'êtes tous venu en aide, à un moment ou à un autre, pour quelques sujets que ce soit. Alors merci à tous, car le fait de vous côtoyer, de près comme de loin, à participer à mon bien-être en VIPAC et à forger mon comportement professionnel. J'espère seulement m'être également montré accueillant et aidant à votre égard. Merci pour ces trois ans et bisous à vous tous !

Merci à Evelyne pour m'avoir accueilli dans son bureau durant ma dernière année, pour les conversations, pour ton écoute et tes conseils.

Merci à Chantal et à Céline, pour avoir assuré ma formation technique, à la paillasse comme en expérimentation animale. Vous m'avez enseigné, vous m'avez aidé, et je ne peux que vous en remercier. Travailler à vos côtés fût pour moi un véritable plaisir.

Merci à Rachel, Marion, Claire, Lili, Céline, Annonciade, Florent, François-Xavier, Clément, et les autres doctorants de l'ANSES, pour les apéros, les discussions... Merci à vous tous, pour tous les à-côtés du boulot que nous avons pu faire ensemble. Vraiment dommage de ne pas avoir pu en faire plus, mais que cela ne signifie pas que nous ne devons plus en faire.

Merci à Céline, pour m'avoir fait découvrir le squash, et avoir tenté (vainement) de m'initier à ce sport. J'en garderai le souvenir d'une intense douleur. Merci à toi !

Merci à mes amis, à mes parents et à mes frères. Depuis le début vous m'avez encouragé, dans les études comme dans la vie. J'ai conscience de ne pas toujours avoir été facile, mais visiblement vous n'avez pas trop l'air de m'en tenir rigueur. J'imagine que c'est ça une famille. Merci à vous, pour absolument tout.

Merci à tous

Béatrice... Paul... je ne peux pas vous dire merci. Je pense sincèrement que ce mot est trop faible pour refléter ma pensée à votre égard.

Quand je t'ai rencontré Béatrice, aux journées francophones de virologie en 2017, j'étais arrivé au stade où je voulais plaquer mes études et mes projets professionnels. Te parler ce jour-là a été la meilleure chose qui pouvait m'arriver. Par la suite, tu as toujours su trouver les mots pour étayer mes réflexions et atténuer mes doutes. Je te dois énormément.

Paul, tu as été une source d'apprentissage extraordinaire. Les moments passés ensemble à corriger les articles ont été super. Pour moi, pouvoir profiter de ton enseignement rédactionnel a été ma plus grande joie, et tes conseils résonneront longtemps dans mon esprit.

Je vous dois beaucoup, à tous les deux. Durant ces trois années, vous m'avez été d'un très grand soutien, vous m'avez enseigné beaucoup plus de choses que je ne pouvais espérer en apprendre durant un doctorat, sur les plans professionnel, scientifique et humain et nos réunions ont souvent été un moment de rigolade très apprécié. J'ai pour vous une très grande affection et, à défaut de peut-être ou pas travailler à nouveau ensemble, j'espère sincèrement que nous provoquerons souvent les occasions de nous revoir.

Emeline... Cela fait maintenant 10 ans que nous sommes ensemble. Nous avons eu des hauts et des bas, comme beaucoup de couples j'imagine, mais au final, ce qui importe c'est que je veuille passer ma vie avec toi. Alors oui, ce paragraphe est un peu court, mais nous avons du temps pour le construire. Ensemble. Tu es ma meilleure amie et je suis fier d'être devenu ton mari.

Pour les initiés...



Table des matières

Liste des abréviations	11
Liste des Figures	13
Liste des Tableaux	15
1 Brève histoire de la découverte des Coronavirus	19
2 Description des <i>Gammacoronavirus</i> aviaires	22
2.1 Structure, morphologie et caractéristiques physico-chimiques du virion	22
2.2 Organisation génomique.....	23
2.3 Protéines des <i>Gammacoronavirus</i> aviaires	23
2.3.1 Les protéines non structurales	23
2.3.2 Les protéines structurales	25
2.3.3 Les protéines accessoires.....	26
2.3.4 Focus sur la protéine S des <i>Gammacoronavirus</i> aviaires.....	27
2.4 Cycle viral des <i>Gammacoronavirus</i> aviaires.....	28
2.4.1 Attachement du virus à la surface cellulaire.....	29
2.4.2 Entrée du virus dans la cellule.....	30
2.4.3 Traduction et Transcription/Réplication du génome viral.....	31
2.4.4 Assemblage/Bourgeonnement.....	33
2.5 Méthodes de culture des coronavirus aviaires.....	34
2.5.1 Propagation sur culture cellulaire	34
2.5.2 Propagation sur explants.....	35
2.5.3 Propagation sur œufs	35
3 Maladies et épidémiologie des <i>Gammacoronavirus</i> aviaires	37
3.1 La bronchite infectieuse causée par l'IBV	37
3.1.1 Maladies, espèces infectées et tissus cibles	37
3.1.2 Classification des souches et épidémiologie	39
3.2 Coronavirose de la dinde causée par le TCoV	41
3.2.1 Maladie, espèce infectée et tissus cibles.....	41
3.2.2 Classification des souches et épidémiologie	42

3.3	Coronavirose de la pintade causée par le GFCoV	43
3.4	Les <i>Gammacoronavirus</i> dans d'autres espèces d'oiseaux	43
3.5	La vaccination contre les <i>Gammacoronavirus</i> aviaires	44
3.5.1	Les différents types de vaccins	44
3.5.2	Conséquences de certaines mauvaises pratiques vaccinales	45
3.6	Les méthodes de détection des coronavirus aviaires	45
3.6.1	Méthodes de détection immunologique <i>in vitro</i>	45
3.6.2	Méthodes de détection moléculaire	46
3.6.3	Méthode de détection par Séquençage Haut-Débit	47
4	Evolution des génomes viraux	49
4.1	Les théories de l'évolution	49
4.2	Création et transmission de la diversité	49
4.3	Le système immunitaire adaptatif comme pression de sélection	51
4.4	Analyse de l'évolution des génomes	53
5	Problématique de la thèse	55
6	Développement d'un outil de visualisation des variants génétiques viraux	59
7	Enrichissement des échantillons en ARN viral	71
8	Evolution des <i>Gammacoronavirus</i> aviaires TCoV et IBV en absence et en présence d'immunité vaccinale préexistante	87
8.1	Evolution du TCoV en absence d'immunité vaccinale pré-existante	87
8.2	Evolution de l'IBV en absence ou présence d'une immunité vaccinale pré- existante	101
9	Discussion Générale	127
9.1	Améliorer la détection des variants génétiques	127
9.2	L'évolution en l'absence d'immunité préexistante	128
9.3	L'évolution en présence d'une immunité préexistante	130
10	Perspectives	132
11	Références bibliographiques	135

Liste des Abréviations

ADN : Acide désoxyribonucléique

ADNc : Acide désoxyribonucléique complémentaire

ARN : Acide ribonucléique

COVID-19 : *Coronavirus Disease 2019*

ELISA : *Enzyme Link ImmunoSorbent Assay*

EOPS : Exempt d'Organismes Pathogènes Spécifiés

ER : *Endoplasmic Reticulum* – Réticulum endoplasmique

ERGIC : *Endoplasmic Reticulum Golgi Intermediate Compartment* – Compartiment intermédiaire entre ER et l'appareil de Golgi

GFCoV : *Guineafowl Coronavirus* – Coronavirus de la pintade

HVR : *Hypervariable Region* – Région hypervariable

IBV : *Infectious Bronchitis Virus* – Virus de la bronchite infectieuse

NGS : *Next Generation Sequencing* – Deuxième génération de séquençage, aussi appelé Séquençage haut débit

NSP : *Non Structural Protein* – Protéine non structurale

NTD/CTD : *N/C-Terminal Domain* – Domaine N/C-terminal

ORF : *Open Reading Frame* – Cadre de lecture ouvert

PCR : *polymerase chain reaction* - Réaction de polymérase en chaîne

qRT-PCR : *quantitative reverse transcription PCR* - PCR quantitative après transcription inverse

SARS-CoV2 : *Severe Acute Respiratory Syndrom Coronavirus 2*

sgARN : ARN sous génomique

TCoV : *Turkey Coronavirus* – Coronavirus de la dinde

TRS-L : *Transcription Regulatory Sequence Leader* – Séquence leader de régulation de la transcription

TRS-b : TRS-body

UTR : *Untranslated Region* - Région non traduite

Liste des Figures

Figure 1: Différents coronavirus observés en microscopie électronique à transmission.....	19
Figure 2: Taxonomie des coronavirus selon le Comité International de Taxonomie des Virus.....	20
Figure 3: Schéma de l'organisation structurale de la particule virale des Gammacoronavirus aviaires et de leur génome.	22
Figure 4: Structure de l'ectodomaine de la protéine de spicule de l'IBV (souche IBV-Beaudette) déterminée par cryoélectro microscopie.....	27
Figure 5: Représentation schématique du cycle viral.....	28
Figure 6: Description d'un œuf embryonné pour inoculation de coronavirus aviaires.	36
Figure 7 : Classification génotypique des souches IBV.....	39
Figure 8 : Répartition géographique mondiale des différents géotypes IBV.....	40
Figure 9: Processus de détermination de séquences par séquençage haut-débit... ..	48
Figure 10 : Illustration d'une sélection de population virale par l'intermédiaire des anticorps neutralisants.	52
Figure 11: Idéogramme décrivant le processus de traitement des données par l'outil Viral Variant Visualiser.	60
Figure 12: Représentation schématique de la notion de profondeur de couverture.	62
Figure 13: Comparaison des profondeurs de couvertures issues du séquençage IBV, avant et après inoculation sur œuf de poule embryonné.....	72
Figure 14: Comparaison des populations virales avant et après un passage sur œuf de poules.....	73

Liste des Tableaux

Tableau 1 : Liste des protéines non structurales des <i>Coronaviridae</i> et leurs rôles associés.	24
Tableau 2 : Variants génétiques identifiés après sélection au cours des passages sur dindes lors de l'expérimentation Fr-TCoV	89
Tableau 3: Variants génétiques identifiés après sélection au cours des passages sur poulets lors de l'expérimentation IBV hétérologue H120-D388	103

Première partie
-
Introduction

1 Brève histoire de la découverte des Coronavirus

Au début des années 1930, une nouvelle maladie respiratoire a été décrite dans un élevage de poulets situé dans le Dakota du Nord aux Etats-Unis d'Amérique (USA). Cette maladie est alors nommée la bronchite infectieuse aviaire. Deux ans plus tard, des scientifiques montrent que le pathogène responsable de cette nouvelle maladie est un agent ultra filtrable (virus) qui remplit le postulat de Koch¹. Ils le nomment alors virus de la bronchite infectieuse (IBV) (Bushnell and Brandly, 1933; Fabricant, 1998; Schalk and Hawn, 1931). Au cours des 40 années qui suivent, des virus ressemblant à l'IBV ont été décrits dans différentes espèces animales comme les souris (*Mouse Hepatitis Virus*), les rats (*Rat Coronavirus, Sialodacryoadenitis Virus*), les porcs (virus de la gastroentérite transmissible, virus de l'encéphalomyélite hémagglutinante), mais aussi l'Homme (B814, 229E). Ce n'est qu'à la fin des années 1960 qu'un consortium de huit virologues, mettant en commun leurs observations, conclut que les différents isolements viraux observés en microscopie électronique correspondent à une même famille virale. Ces scientifiques prennent alors la décision de nommer ce nouveau groupe Coronavirus, du fait de leur apparence en couronne (*Corona* en latin) observée sur les images de microscopie (McIntosh, 1974; Tyrrell et al., 1968) (Figure 1).

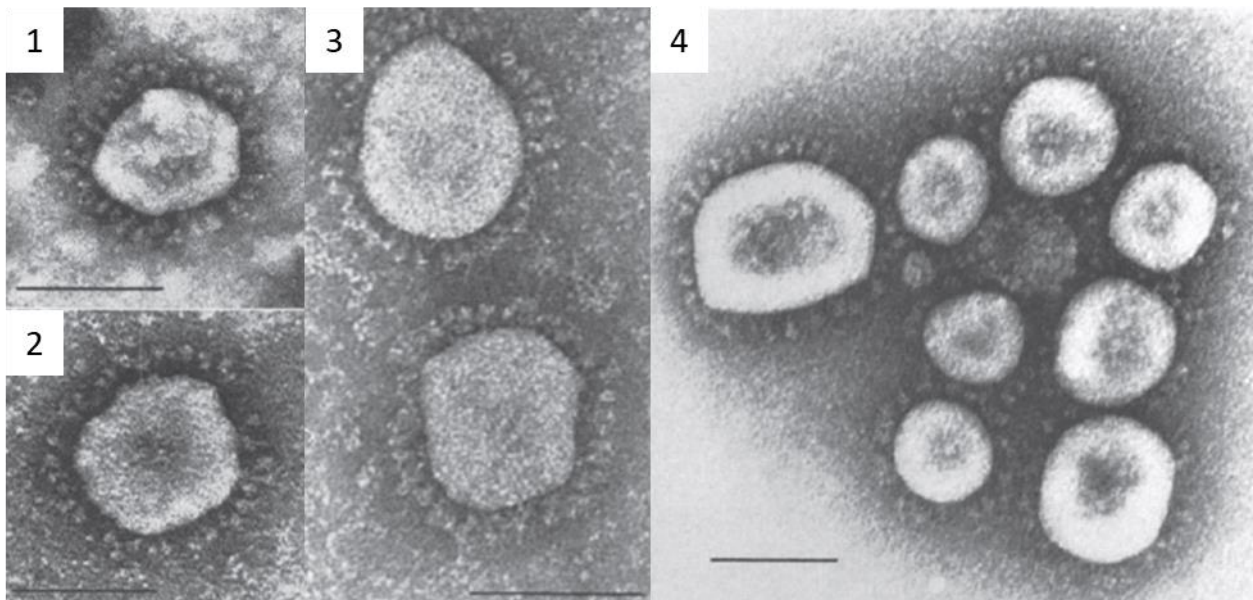


Figure 1: Différents coronavirus observés en microscopie électronique à transmission. 1) Coronavirus Humain, souche 229E ; 2) Coronavirus Humain, souche OC16 ; 3) Virus de la bronchite infectieuse, souche Beaudette ; 4) Virus de l'hépatite murine, souche A59. Les barres représentent une échelle de 100 nm. Images issues de McIntosh (1974).

¹ Postulat de Koch : le micro-organisme doit être isolé à partir d'un sujet malade ; le micro-organisme doit être cultivable et isolable ; le micro-organisme isolé doit pouvoir reproduire la maladie après inoculation à de nouveaux sujets sensibles ; le micro-organisme doit pouvoir être ré-isolé à partir des sujets inoculés.

Depuis, le groupe des Coronavirus n'a cessé de s'agrandir avec la détection et l'isolement continuel de nouvelles souches chez différentes espèces, le dernier en date étant le SARS-CoV2, virus responsable de la pandémie de COVID-19. Le groupe des Coronavirus a été organisé en une grande famille taxonomique appelé *Coronaviridae*, qui a été officiellement reconnue en 1975 par le Comité International de Taxonomie des Virus (ICTV - https://talk.ictvonline.org/taxonomy/p/taxonomy_releases). Cette famille appartient au règne *Riboviria*, au royaume des *Orthornavirae*, au phylum des *Pisuviricota*, à la classe des *Pisoniviricetes*, à l'ordre des *Nidovirales*, et au sous-ordre des *Cornidovirineae*. Actuellement, la famille des *Coronaviridae* se décompose en deux sous-familles : *Letovirinae*, qui inclut un seul genre viral *Alphaletovirus* ; et *Orthocoronavirinae* qui inclut quatre genres viraux *Alphacoronavirus*, *Betacoronavirus*, *Gammacoronavirus* et *Deltacoronavirus* (ICTV - Figure 2).

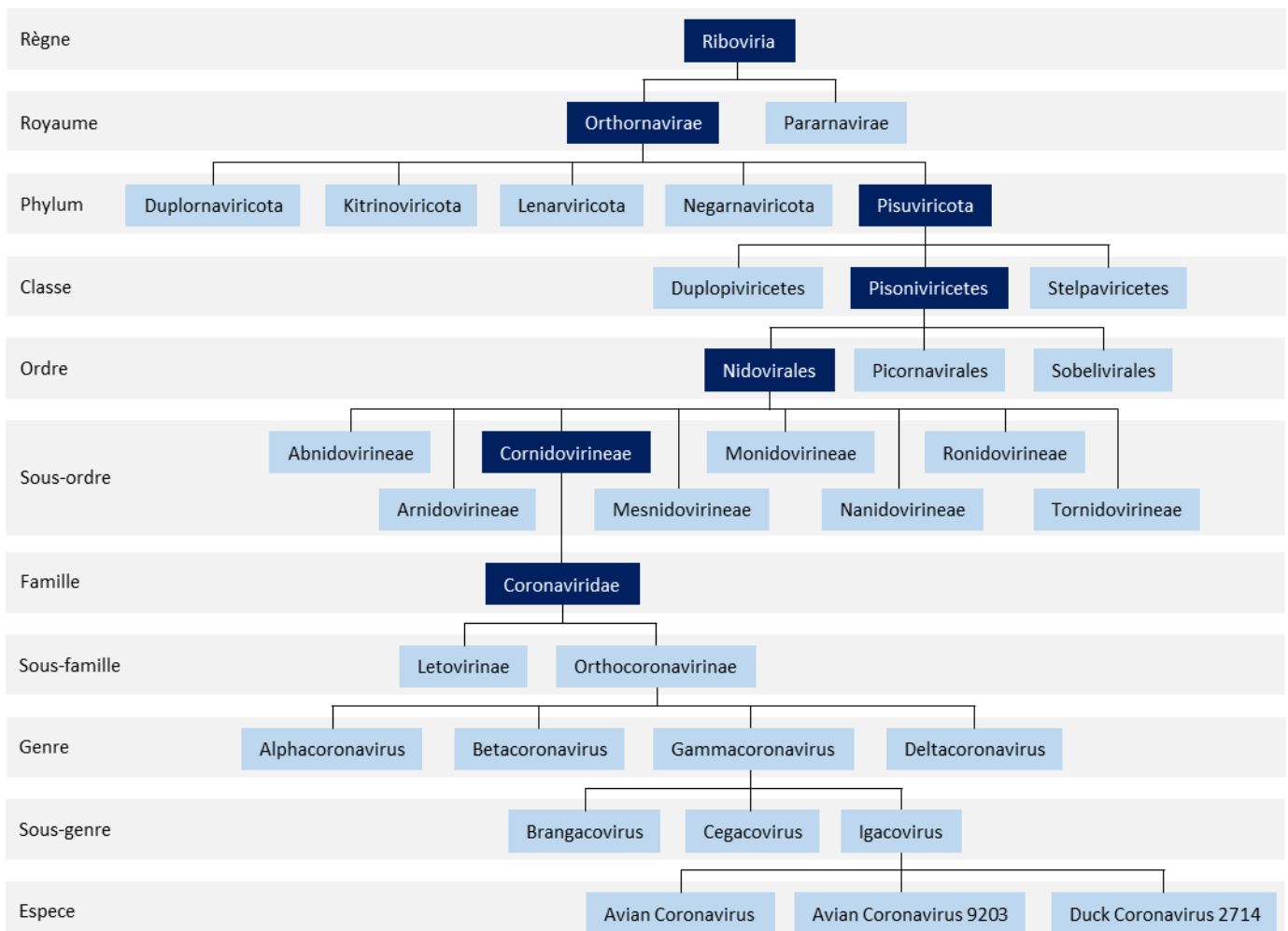


Figure 2: Taxonomie des coronavirus selon le Comité International de Taxonomie des Virus. Mise à jour effectuée le 22 mai 2020.

Bien que la sous-famille *Letovirinae* et le genre *Alphaletovirus* soient relativement récents (Juillet 2018 avec la découverte d'un génome de *Microhyla letovirus*, un virus de batracien), les quatre genres appartenant aux *Orthocoronavirinae* sont les genres « historiques » des Coronavirus. Les *Alpha-* *Beta-* et *Gammacoronavirus*, dont les noms sont apparus en 2008 et officialisés en 2011 (King et al., 2011), étaient initialement dénommés en tant que Coronavirus des groupes I, II et III respectivement. Le genre des *Deltacoronavirus* n'a officiellement été créé qu'en 2011 (ICTV), à la suite de la découverte de nouveaux génomes de Coronavirus (Woo et al., 2009). Pour ces quatre genres viraux, 25 sous-genres se répartissent de la façon suivante : 14, 5, 3, et 3 sous-genres respectivement. Sur la base de comparaisons génétiques et de la classification taxonomique, il semblerait que le plus lointain ancêtre des Coronavirus soit apparu chez les animaux volants, et qu'à la suite de différents échanges entre les oiseaux et les chauve-souris, une spéciation soit survenue (Woo et al., 2012). À partir de ce moment, les chauves-souris seraient à l'origine des *Alphacoronavirus* et des *Betacoronavirus* (comme le virus de la diarrhée épidémique porcine et le virus hémagglutinant de l'encéphalomyélite porcine respectivement), alors que les oiseaux seraient à l'origine des *Gammacoronavirus* et des *Deltacoronavirus* (comme le virus de la bronchite infectieuse et le bulbul coronavirus respectivement).

Actuellement, à l'exception de génomes de *Gammacoronavirus* détectés chez la baleine blanche (Mihindukulasuriya et al., 2008), et chez les dauphins (Woo et al., 2014) et de *Deltacoronavirus* détecté et isolé chez des porcs (Hu et al., 2015; Woo et al., 2012), les virus de ces deux genres viraux infectent des oiseaux. Jusqu'à aujourd'hui, les *Deltacoronavirus* aviaires ont été détectés chez des oiseaux sauvages, et les *Gammacoronavirus* aviaires ont été retrouvés chez les oiseaux d'élevage. Deux *Gammacoronavirus* aviaires, du sous-genre des *Igacovirus* ont fait l'objet d'étude du présent projet de thèse : l'IBV et le Coronavirus de la dinde (TCoV).

2 Description des *Gammacoronavirus* aviaires

2.1 Structure, morphologie et caractéristiques physico-chimiques du virion

La particule virale présente un aspect pléomorphe, le plus souvent sphérique (Figure 1), avec un diamètre qui peut varier entre 80 et 160 nm (Berry et al., 1964; McIntosh, 1974; Sturman and Holmes, 1983). Les particules virales ont des projections qui partent vers l'extérieur et qui leur confèrent leur aspect dit « en couronne ». Les coronavirus aviaires présentent six éléments majeurs, communs à tous les membres des *Coronaviridae*. Une bicouche lipidique compose la membrane du virus et permet l'ancrage de trois protéines virales nommées protéine de Spicule (S), protéine d'Enveloppe (E), protéine de Membrane (M). À l'intérieur de la particule virale se trouve un génome ARN et une quatrième protéine de structure, la Nucléoprotéine (N), qui recouvre le génome viral afin de former une nucléocapside à symétrie hélicoïdale. L'ensemble de la particule virale a une densité de 1,18 g/mL (Macnaughton and Davies, 1980) (Figure 3A).

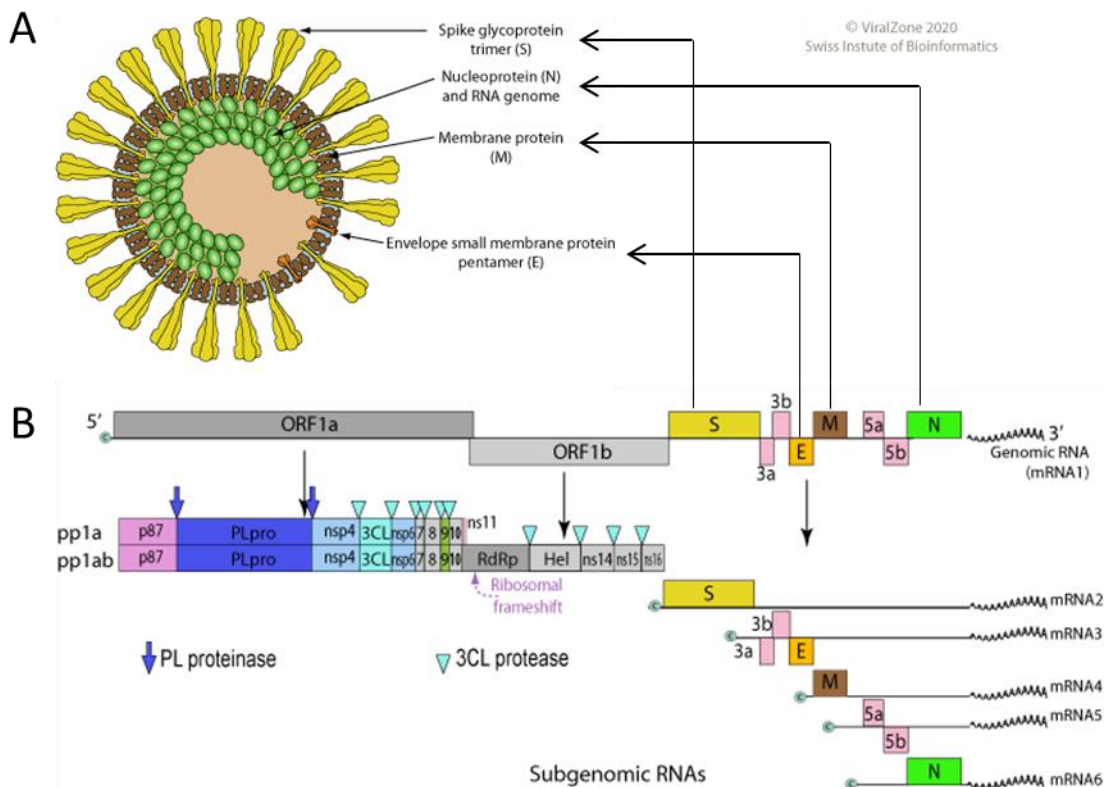


Figure 3: Schéma de l'organisation structurale de la particule virale des *Gammacoronavirus* aviaires et de leur génome. A) Représentation de la particule virale, avec une membrane dans laquelle est retrouvée les protéines de structure du virus : la protéine de membrane M (marron), la protéine de spicule S (jaune), la protéine d'enveloppe E (orange) et le complexe ribonucléoprotéique, formé par l'association de la nucléoprotéine N (vert) et du génome viral. B) Représentation schématique de l'organisation du génome viral, un ARN simple brin de polarité positive. Sont représentés également, les produits de la traduction de l'orf1ab (PP1a et PP1ab) ainsi que les ARN viraux sous-génomiques produits à la suite d'une transcription discontinue du génome viral (trois ARN monocistroniques et deux ARN polycistroniques). Image provenant de ViralZone, Swiss Institute of Bioinformatics).

2.2 Organisation génomique

Les génomes de coronavirus font partie des plus grands génomes ARN (Watkins et al., 1975) actuellement connus, avec une taille moyenne de 27,5 kb dans le cas des *Gammacoronavirus* aviaires (Figure 3B). Ces génomes sont constitués d'un ARN simple brin de polarité positive (MacNaughton, 1978), c'est-à-dire directement traduisible par les ribosomes cellulaires après l'entrée du génome dans la cellule. La toute première séquence complète d'un génome de coronavirus a été établie pour l'IBV (Bournsell et al., 1987) et bien que les génomes de chacune des souches de *Coronaviridae* possèdent leurs propres particularités, cette première séquence a révélé l'organisation génomique partagée par tous les membres de cette famille virale.

À l'instar des ARN messagers cellulaires, l'extrémité 5' des génomes de *Gammacoronavirus* est coiffée par l'ajout d'une 7-méthylguanosine qui précède une région non traduite dite 5'UTR (*Untranslated Region*) (Hulo et al., 2011). L'extrémité 3' des génomes est formée d'une région non traduite dite 3'UTR à la fin de laquelle se trouve une queue poly-A (MacNaughton, 1978; Schochetman et al., 1977) (Figure 3B). La coiffe et la queue poly-A sont des marqueurs d'ARN messagers cellulaires qui permettent au génome viral de leurrer la défense intrinsèque de la cellule et ainsi échapper à la dégradation par des facteurs de restriction cytosolique (Hyde and Diamond, 2015; Linares-Fernández et al., 2020). Ces marqueurs permettent également l'initiation de la traduction de l'ARN génomique viral dès son entrée dans la cellule (Nakagawa et al., 2016). Au sein des régions 5' et 3' UTR se trouvent différentes structures secondaires de l'ARN impliquées dans la régulation de la transcription/réplication du génome viral (Dalton et al., 2001; Yang and Leibowitz, 2015). Situés entre les deux UTRs, cinq gènes essentiels, nommés orf1ab, S, E, M et N, sont retrouvés dans le génome de tous les coronavirus. Occupant près de deux tiers du génome viral et long de plus de 20 kb, l'orf1ab code pour des protéines impliquées dans le cycle viral (réplication, transcription) et dans la mise en place d'organites de réplication. Dans le dernier tiers du génome viral se trouvent les gènes S, E, M, N qui codent pour quatre protéines de structures. Dans cette même région sont également retrouvés différents gènes accessoires, dont le nombre varie en fonction des souches. Ces gènes sont dits accessoires car l'activité des protéines encodées n'est pas nécessaire à la réplication virale *in vitro*, et leur absence *in vivo* entraîne une réduction de l'infectivité du virus dans l'hôte.

2.3 Protéines des *Gammacoronavirus* aviaires

2.3.1 Les protéines non structurales

L'orf1a/1ab représente les deux tiers proximaux du génome viral, soit près de 20kb.

Tableau 1 : Liste des protéines non structurales des *Coronaviridae* et leurs rôles associés.

Nom	Fonction de la protéine	Références
NSP1	Absente des <i>Gammacoronavirus</i> , serait impliquée dans la réduction de la traduction des ARNm cellulaires	(Narayanan et al., 2015)
NSP2	Fonction inconnue, n'est pas forcément nécessaire à la réplication	(Graham et al., 2006)
NSP3	Protéase Papaine-like : clivage entre différentes NSP Rôle dans la formation des organites de réplication	(Lei et al., 2018) (Doyle et al., 2018)
NSP4	Rôle dans la formation des organites de réplication	(Doyle et al., 2018)
NSP5	Cystine protéase : clivage entre différentes NSP	(Fields et al., 2013)
NSP6	Détourne le mécanisme d'autophagie Rôle dans la formation des organites de réplication	(Cottam et al., 2014, 2011; Doyle et al., 2018)
NSP7	Avec NSP8 : forme le cofacteur de processivité de NSP12	(Subissi et al., 2014)
NSP8	Initialement décrite comme une primase permettant de synthétiser une amorce initiatrice de transcription du génome, plus récemment décrite comme étant l'acteur qui permet de synthétiser la queue polyA des ARN viraux Avec NSP7 : forme le cofacteur de processivité de NSP12	(Imbert et al., 2006) (Tvarogová et al., 2019) (Subissi et al., 2014)
NSP9	Fixation à ARN simple brin potentiellement pour fixer le complexe de réplication	(Zeng et al., 2018)
NSP10	Cofacteur de NSP14 et NSP16	(Bouvet et al., 2014)
NSP11	Aucun rôle de décrit Extrémité C-terminale de PP1a, taille de 23 acides aminés Les 15 premiers acides aminés correspondent aux 15 premiers acides aminés de NSP12	/
NSP12	ARN polymérase ARN dépendante	(Tan et al., 2018)
NSP13	Hélicase : permet d'ouvrir les structures ARN du génome viral Activité phosphatase : retire le phosphate gamma du GTP ajouté afin de former la coiffe	(Ivanov and Ziebuhr, 2004)
NSP14	3'5' exoribonucléase : confère au complexe de transcription/réplication une activité de relecture Ajoute le méthyl en position 7 de la base G qui sert de coiffe aux ARN viraux	(Eckerle et al., 2007) (Jin et al., 2013)
NSP15	Endoribonucléase : clive les ARN après des uridines Fonction associée à la digestion de la séquence polyU présente à l'extrémité 5 des ARN viraux antisens et qui semble être un marqueur de danger pour la cellule, détectée par MDA5	(Cao et al., 2008a) (Hackbart et al., 2020)
NSP16	2'O-méthyltransférase : ajoute un groupement méthyl sur le premier ribose de ARN en 2' lors de la formation de la coiffe	(Decroly et al., 2008)

Les orf1a et 1ab codent pour deux polyprotéines (PP), nommées PP1a et PP1ab, qui sont produites dans des proportions respectives d'environ 70 et 30% (Brierley et al., 1987). Lors de la production de PP1a, une structure particulière de l'ARN associée à une séquence dite glissante (localisées au milieu de l'orf1a/1ab) permet au ribosome de reculer d'un nucléotide, décalant ainsi le cadre de lecture, pour permettre la production de PP1ab (Brierley et al., 1992, 1991).

Afin de permettre la libération de toutes les protéines non structurales (NSP), les deux polyprotéines sont séquentiellement clivées par l'action combinée de deux protéases virales (nommées NSP3 et NSP5) présentes au sein de PP1a. Après clivage, quinze NSP sont produites (NSP2 à 11, produites par PP1a, NSP2 à 16 produites par PP1ab). Une seizième protéine, NSP1, est retrouvée chez les *Alphacoronavirus* et *Betacoronavirus*, mais n'est pas présente chez les *Gammacoronavirus*. Les NSP sont des protéines multifonctionnelles impliquées dans divers mécanismes comme la diminution de la traduction des ARNm cellulaires, la modulation de la défense cellulaire intrinsèque vis-à-vis de l'infection, la formation d'organites de réplication, mais également la production d'un complexe multi-protéique de transcription/réplication du génome viral, formé par l'association des NSP7, 8, 9, 10, 11/12, 13 et 14 (Snijder et al., 2016).

2.3.2 Les protéines structurales

Le gène S code pour la protéine de spicule (aussi appelé protéine S). Cette protéine est le principal déterminant antigénique du virus et constitue la partie la plus exposée du virus (Figure 3, en jaune). Pour être fonctionnelle, cette protéine s'auto-assemble en trimère. Chaque monomère est clivé en deux sous-domaines (S1 et S2), qui restent liés entre eux de façon non covalente. Dans sa partie N-terminale (S1), le trimère présente le site de fixation aux récepteurs cellulaires et dont la reconnaissance par des anticorps peut neutraliser le virus, l'empêchant d'entrer dans la cellule. Malgré la présence du site de fixation au récepteur cellulaire, S1 est la région la plus variable de la protéine, permettant ainsi au virus d'échapper aux anticorps neutralisants. Dans sa partie C-terminale (S2) se trouve le peptide de fusion qui permet la fusion de la membrane virale avec celle de la cellule lors de l'entrée virale (Wickramasinghe et al., 2014).

Le gène E, aussi appelé gène 3c chez les *Gammacoronavirus* aviaires, code pour la protéine d'enveloppe (Figure 3, en orange), qui est la protéine de structure la moins représentée de la particule virale. Cette protéine se pentamérise pour former une viroporine, un canal ionique localisé dans la membrane de l'appareil de Golgi et où elle permet de modifier son pH luminal probablement afin de préserver l'intégrité de la protéine de spicule lors de son passage dans l'appareil de Golgi (Westerbeck and Machamer, 2015, 2019).

Le gène M code pour la protéine la plus abondante de la particule virale, la protéine de membrane (Figure 3, en marron), qui donne sa forme à la particule. Cette protéine possède trois domaines transmembranaires, un petit domaine N-terminal extravirion et un domaine C-terminal intravirion plus important. Dans les membranes, il semblerait que la protéine M se dimérise et est alors capable de prendre deux formes : la première, dite compacte, permettrait de fixer le complexe ribonucléoprotéique formé par l'association du génome viral et de la nucléoprotéine ; la seconde, dite longue, joue son rôle au moment du bourgeonnement afin de rigidifier la membrane de la particule virale (Neuman et al., 2011).

Le gène N code pour la nucléoprotéine (Figure 3, en vert), qui est fixée au génome viral mais qui ne permet qu'une protection partielle du génome viral contre une digestion enzymatique par les ribonucléases (Davies et al., 1981; Macneughton and Davies, 1978). La nucléoprotéine est phosphorylée (Jayaram et al., 2005) et se compose de trois domaines. Les domaines dits NTD (*N-terminal domain*) et CTD (*C-terminal domain*) fixent le génome viral, tandis que le troisième, nommé Domaine 3, interagit avec la protéine M, probablement afin de permettre l'empaquetage du génome viral lors du bourgeonnement de la particule virale (McBride et al., 2014).

2.3.3 Les protéines accessoires

Les génomes des *Coronaviridae* possèdent également des gènes accessoires dont le nombre, la localisation sur le génome et la dénomination varient en fonction du type viral. Dans le cas des coronavirus aviaires IBV et TCoV quatre gènes accessoires ont été identifiés. Les gènes accessoires 3a et 3b sont situés entre les gènes S et E et les protéines encodées impactent la défense antivirale intrinsèque de la cellule en modulant la réponse interféron (Kint et al., 2015). Les gènes accessoires 5a et 5b sont situés entre les gènes M et N mais le rôle des protéines encodées est encore méconnu. Dans le cas du TCoV, quatre autres cadres de lectures ont été également prédits : les gènes 4b et 4c situés en aval du gène M, le gène 6b situé en aval du gène N et l'orf X situé entre les gènes M et 5a (Brown et al., 2016; Cao et al., 2008b). Cependant, aucune preuve n'a été apportée quant à l'expression des gènes 4b, 4c, 6b et orf X *in vitro* ou *in vivo* (Gomaa et al., 2008). À l'heure actuelle, et cela est valable pour l'ensemble des membres de *Coronaviridae*, aucun gène accessoire n'a encore été décrit entre les gènes E et M.

Ces gènes sont dit accessoires car les virus ne les possédant pas (ou possédant des gènes accessoires non fonctionnels) peuvent se répliquer *in vitro*, et *in vivo*, l'absence de ces gènes permet une atténuation du virus (Laconi et al., 2018; van Beurden et al., 2018). Par ailleurs, il existe des souches naturelles d'IBV infectieuses qui présentent une organisation génomique dépourvue de gènes accessoires (Hewson et al., 2011).

2.3.4 Focus sur la protéine S des *Gammacoronavirus* aviaires

La protéine S est une protéine à un seul domaine transmembranaire dont l'extrémité N-terminale est située en extracellulaire (protéine transmembranaire de type I). Cette protéine est également une protéine de fusion de type I dont le fonctionnement nécessite une trimérisation et un clivage (Bosch et al., 2003). Du fait de la grande diversité de séquences, la taille de la protéine S est variable, actuellement comprise entre 1140 et 1230 acides aminés. En absence de modifications post-traductionnelles, la masse moléculaire de cette protéine est d'environ 128 kDa, mais peut atteindre 200 kDa grâce aux multiples sites potentiels de N-glycosylation (Shang et al., 2018; Wickramasinghe et al., 2014) qui permettraient de moduler la fixation de la protéine sur les acides sialiques (Bouwman et al., 2020). La protéine S se décompose en trois segments: un ectodomaine, un domaine transmembranaire et un domaine intravirion. La présence d'un site de clivage (motif RRFRR/S) dans l'ectodomaine (Cavanagh et al., 1986) permet la dissociation de deux sous unités, S1 (en N-terminal) et S2 (en C-terminal). Ce clivage est nécessaire à l'activité de fusion de la protéine et est réalisé par une protéase cellulaire de type furine.

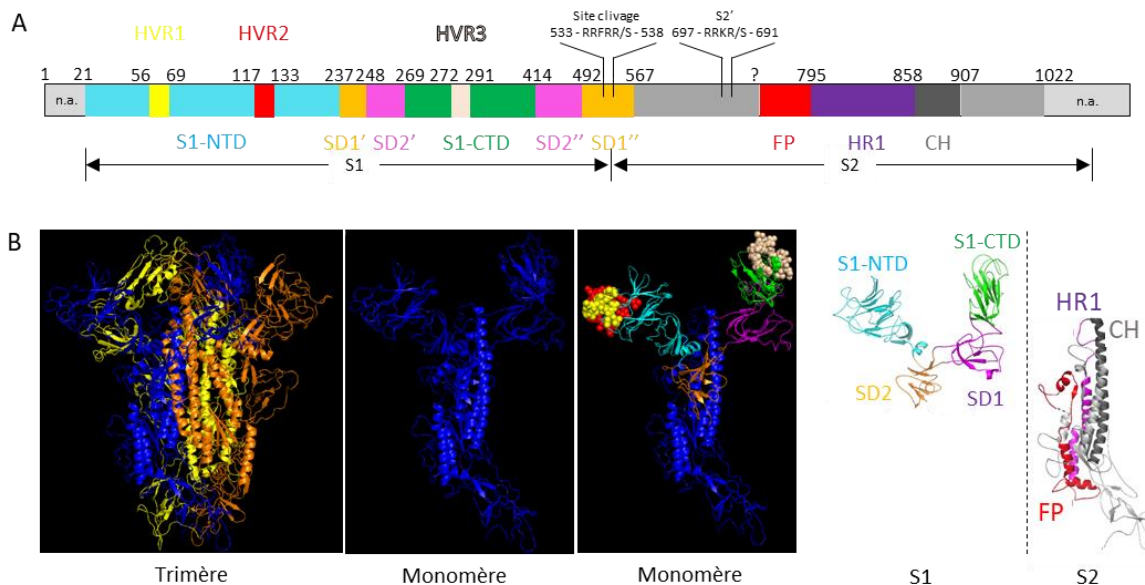


Figure 4: Structure de l'ectodomaine de la protéine de spicule de l'IBV (souche IBV-Beaudette) déterminée par cryoélectro microscopie. A) représentation linéaire des différentes régions de l'ectodomaine de la protéine S. B) représentation tridimensionnelle de la forme préfusion du trimère et du monomère S à la surface de la particule virale, les régions en sphère dans l'image du deuxième monomère représentent les régions hypervariables (HVR). N/CTD : domaine N ou C-terminal, SD : sous domaine, FP : peptide de fusion, HR : région répétée en heptade, CH : hélice centrale, n.a. : non affichée. Les images ont été modifiées à partir de Shang et al., 2018.

La sous-unité S1 est impliquée dans la reconnaissance du récepteur (Hulswit et al., 2016). Chez l'IBV, il a été montré que cette sous-unité est composée de quatre domaines : un domaine N-terminal (S1-NTD) et un domaine C-terminal (S1-CTD) qui fonctionnent comme des domaines de liaison aux récepteurs et deux domaines additionnels SD1 et SD2 formés par le rapprochement de séquences protéiques distantes (SD1' avec SD1'' et SD2' avec SD2'') dont les rôles sont encore

inconnus (Shang et al., 2018). S1 est également caractérisée par la présence de trois régions hypervariables (HVR1, 2 et 3) (Figure 4) (Kusters et al., 1989; Shan et al., 2018) au niveau desquelles des épitopes spécifiques du sérotype viral sont retrouvés. La fixation d'anticorps sur ces régions peut permettre de neutraliser les virus (Cavanagh et al., 1988). La sous-unité S2 est plus conservée entre les différentes souches d'IBV que la sous-unité S1. Elle est impliquée dans le processus de fusion des membranes virales et cellulaires (Hulswit et al., 2016). S2 contient deux régions répétées en heptade (*heptad repeat* - HR1 (Figure 4) et HR2 non-indiquée), une succession d'acides aminés hydrophobes qui compose le peptide de fusion et dont le rôle est de permettre la fusion des membranes virale et cellulaire, ainsi qu'un domaine transmembranaire qui permet l'ancrage de la protéine dans la membrane virale et un domaine intravirion.

2.4 Cycle viral des *Gammacoronavirus* aviaires

Les différentes étapes du cycle viral sont représentées dans la figure 5.

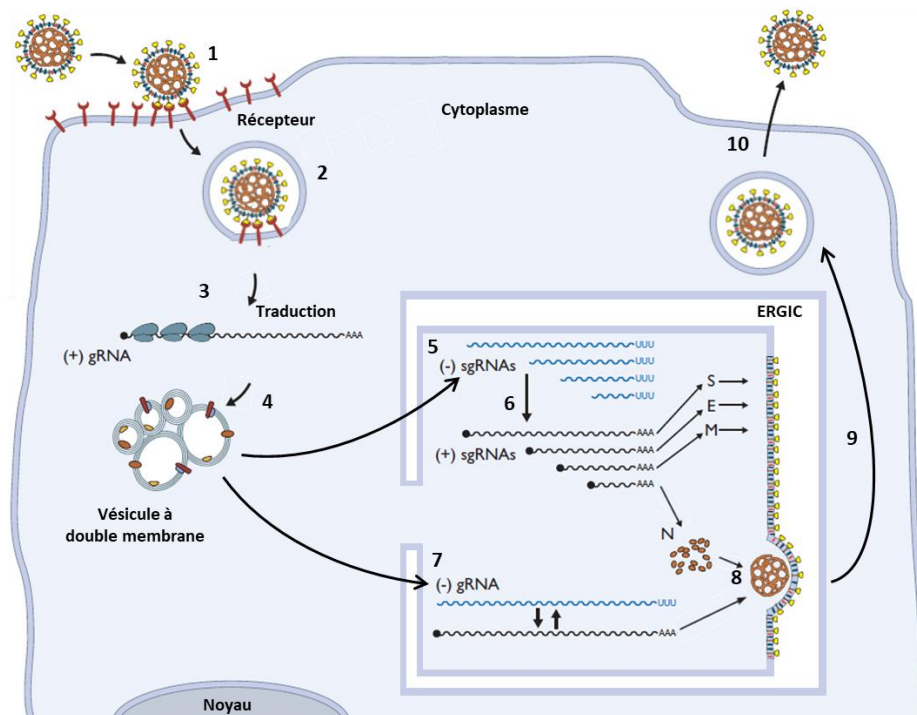


Figure 5: Représentation schématique du cycle viral. Le cycle répliatif du virus est entièrement cytoplasmique. 1) Fixation de la particule virale au récepteur cellulaire ; 2) Endocytose de la particule virale puis fusion de la membrane virale avec celle de l'endosome ; 3) Traduction de l'orf lab pour permettre l'expression des protéines impliquées dans la modulation des fonctions cellulaires et dans la formation du complexe de transcription/réplication ; 4) Formation des organites de réplique dans lesquels se déroule la transcription/réplication du génome viral ; 5) Création des ARN sous-génomiques antisens (-)sgRNA ; 6) Synthèse des ARN sous-génomiques sens (+)sgRNA coiffés et polyadénylés, afin de produire les protéines de structure (S, E M et N) du virus ; 7) Synthèse d'un ARN génomique antisens (-)gRNA puis d'un nouveau génome viral (+)gRNA ; 8) Assemblage et bourgeonnement de la particule virale dans le compartiment intermédiaire entre le réticulum endoplasmique et le Golgi (ERGIC) ; 9) Transport de la vésicule d'exocytose contenant la particule virale vers la membrane plasmique ; 10) Sortie de la particule virale par fusion de la vésicule d'exocytose avec la membrane plasmique. Image modifiée à partir de la figure 28.6 du livre Fields Virology 6^{ème} édition.

2.4.1 Attachement du virus à la surface cellulaire

Afin de permettre l'entrée du virus dans les cellules, celui-ci doit se fixer à un ou plusieurs récepteurs cellulaires. À ce jour, aucun récepteur de nature protéique n'a été décrit comme étant impliqué dans l'entrée des *Gammacoronavirus* IBV ou TCoV dans les cellules cibles. L'aminopeptidase N, un récepteur identifié de certains *Alphacoronavirus* comme le coronavirus de la gastroentérite transmissible du porc ou le coronavirus de la péritonite infectieuse féline (Hulswit et al., 2016; Li, 2015), a été testé comme récepteur potentiel au travers de deux études dont les conclusions sont contradictoires. Une première étude rapporte que l'expression transitoire de l'aminopeptidase N dans des fibroblastes reinaux de hamster (BHK21) permettait l'entrée de IBV-Ark99 dans les cellules, alors même que les BHK21 ne sont pas permissives naturellement à l'IBV (Miguel et al., 2002). La seconde étude contredit la première (Chu et al., 2007).

En revanche, les acides sialiques α -2,3 ont été montrés comme nécessaires pour l'infection de cellules ou d'anneaux de trachée par IBV (Rahman et al., 2009; Winter et al., 2008, 2006) (Figure 5.1). Au cours du cycle viral, la protéine de spicule est clivée en deux sous-unités (S1 et S2) et bien que la sous-unité S1 possède les sites de fixation aux acides sialiques, S2 pourrait également être impliquée dans la fixation des particules virales sur les acides sialiques. En effet, il a été observé que la sous-unité S1 de la souche IBV-Beaudette ne pouvait pas fixer les acides sialiques, alors que l'ectodomaine de la protéine de spicule (correspondant à la partie S1 + S2) le pouvait (Promkuntod et al., 2013). Ce type de récepteur est localisé à la surface des tissus épithéliaux de la trachée (Winter et al., 2008), des poumons (Promkuntod et al., 2014), des intestins (Guo et al., 2007), des reins (Kuchipudi et al., 2009) et des oviductes (Mork et al., 2014), qui sont des tissus infectés par IBV. Cependant, même si des études *in vitro* ont montré que le changement de quelques acides aminés au niveau des régions hypervariables de S1 pouvait engendrer une modification de fixation de S1 sur les acides sialiques, permettant une modulation du tropisme (Bouwman et al., 2020; Saiada et al., 2020), ce type de récepteur ne semble pas être le seul déterminant permettant l'infection des cellules. En effet, il a été montré que les protéines S1 de deux souches IBV néphropathogéniques (IBV-QX et IBV-B1648) pouvaient fixer les acides sialiques présents sur les épithéliums des oviductes, mais que seul la souche IBV-QX était capable d'infecter les cellules épithéliales de ce tissu (Mork et al., 2014). Ainsi, d'autres récepteurs ou facteurs intracellulaires sont probablement impliqués dans la capacité du virus à entrer et se répliquer dans différents tissus.

Un premier exemple est la capacité de la souche IBV-Beaudette à utiliser un autre type de récepteur pour fixer la surface cellulaire. L'IBV-Beaudette est la seule souche à se répliquer

efficacement sur cellules Vero, des cellules d'origine mammifère. Cette capacité à infecter des cellules de mammifère a été supposée possible grâce à l'utilisation des héparanes sulfates (Madu et al., 2007).

Un deuxième exemple est la présence de sites de clivages additionnels au sein de la protéine de spicule, par le changement de quelques acides aminés seulement. Ces sites additionnels semblent être déterminants dans le changement de tropisme des Coronavirus et possiblement dans les cas de sauts d'espèces (Cheng et al., 2019; Yamada and Liu, 2009). La souche IBV-Beaudette possède un deuxième site de clivage (RRRR/S, nommé S2', situé en amont du peptide de fusion) qui pourrait expliquer sa capacité à se répliquer sur cellule Vero et potentiellement d'autres types cellulaires (Yamada and Liu, 2009). Renforçant cette hypothèse, l'introduction artificielle d'un deuxième site de clivage dans la protéine S de la souche IBV-QX a permis à la souche recombinante d'infecter des cellules embryonnaires de poulet, alors que la souche IBV-QX avec un seul site de clivage ne le pouvait pas (Cheng et al., 2019). Par ailleurs, l'inoculation de la souche IBV-QX recombinante à des poulets a révélé que la souche recombinante était capable d'infecter les neurones, alors que la souche avec un seul site de clivage ne le pouvait pas. De plus, les sujets inoculés avec la souche recombinante ont subi des encéphalites ainsi qu'une mortalité prématurée (Cheng et al., 2019). La présence d'un deuxième site de clivage S2' a également été montrée chez plusieurs *Betacoronavirus*, y compris le SARS-CoV2, qui en aurait potentiellement un troisième (Coutard et al., 2020).

Au cours des années 1960, un changement de tropisme d'une souche Mass41 a été observé à la suite d'un refroidissement accidentel de poulets inoculés. Après le refroidissement, la souche respiratoire a changé de tropisme et s'est révélée capable de se répliquer dans les reins et a provoqué des néphropathies (Cook et al., 2012). Cette observation suggère que les conditions environnementales, et donc les conditions d'élevage, influencent le tropisme cellulaire des souches d'IBV.

Dans les cas d'infection par le TCoV, mais également avec les coronavirus de pintade et de caille, les N-glycanes possédant plusieurs répétitions de N-acétyl-lactosamine ont été prédits pour être impliqués dans l'entrée du virus dans les cellules (Wickramasinghe et al., 2015).

2.4.2 Entrée du virus dans la cellule

En utilisant la souche IBV-Beaudette, il a été observé qu'après fixation du virus à la surface de différents types cellulaires (cellules Vero, DF1, Huh7 et H1299), l'internalisation des particules virales dans un endosome est majoritairement effectuée par un mécanisme d'endocytose médié par la clathrine (Wang et al., 2019) (Figure 5.2). Toutefois, lors de l'utilisation des cellules H1299,

l'inhibition de la voie d'endocytose dépendante de la cavéoline a également permis de réduire, à un niveau moindre, l'internalisation du virus, suggérant donc qu'en plus d'un mécanisme majoritaire, un mécanisme d'entrée minoritaire pourrait être utilisé en fonction du type cellulaire (Wang et al., 2019). Après endocytose, l'acidification progressive de l'endosome au cours de sa maturation en lysosome est supposée provoquer des changements de conformation de la protéine S trimérique, permettant le démasquage des peptides de fusion présents dans le cœur hydrophobe de chaque monomère S2. Ces peptides vont alors être propulsés dans la membrane du lysosome afin de permettre la fusion de la membrane virale avec celle du lysosome (Chu et al., 2006). À la suite de cette fusion, le complexe ribonucléoprotéique (ou nucléocapside), formé par l'association du génome viral et des protéines N associées, est libéré dans le cytosol.

2.4.3 Traduction et Transcription/Réplication du génome viral

Après être libéré dans le cytosol, le génome des coronavirus, un ARN coiffé et polyadénylé, est directement traduit par les ribosomes cellulaires afin de produire les polyprotéines PP1a et PP1ab (Figure 5.3). Ces protéines sont clivées par l'action des deux protéases virales (NSP3 et NSP5), permettant ainsi de libérer tous les acteurs impliqués dans la transcription et réplication du génome viral. Certaines de ces protéines provoquent un réarrangement du réseau membranaire de la cellule (Tableau 1), afin de créer des organites (ou sphérules) qui constituent le centre de transcription/réplication du génome viral (Maier et al., 2014) (Figure 5.4). Dans ces centres, l'expression des protéines structurales repose sur la stratégie de transcription/réplication du génome viral. Le complexe de transcription/réplication se positionne sur l'extrémité 3' du génome via la reconnaissance de différentes structures secondaires de l'ARN dans la région 3'UTR. L'initiation de la transcription est encore mal connue. Elle a d'abord été supposée comme débutant par l'action de NSP8 en tant que primase (Imbert et al., 2006), mais cette hypothèse a ensuite été contredite (Subissi et al., 2014; Tvarogová et al., 2019). Après avoir été initiée, NSP12 poursuit le processus de transcription en direction de l'extrémité 5' du génome viral, aidé en arrière par un anneau de processivité formé par l'association de huit NSP7 et huit NSP8 (Subissi et al., 2014; te Velthuis et al., 2012) et aidé en avant par l'hélicase NSP13 qui défait les structures ARN pour permettre le passage du complexe de transcription/réplication. À chaque ajout d'un nouveau ribonucléotide, la validité de cette incorporation semble être vérifiée par NSP14, qui retirerait les nucléotides mal incorporés grâce à son activité 3'-5' exonucléasique (Eckerle et al., 2007).

À un moment de sa progression, le complexe de transcription/réplication rencontre une séquence particulière présente en amont de chaque cadre de lecture, un signal de régulation de la transcription, appelé TRS-b (*Transcription Regulatory Sequence - body*). En rencontrant cette

séquence, le complexe de transcription/réplication a la possibilité de « sauter » directement vers le TRS le plus proche de l'extrémité 5' du génome, appelé TRS-L (*Transcription Regulatory Sequence - Leader*) pour ensuite continuer de transcrire l'extrémité 5' du génome, ou de continuer la transcription du génome viral en direction de l'extrémité 5'. Si le complexe de réplication « saute », alors l'ARN viral créé est plus court que le génome viral, l'ARN viral est alors dit « sous-génomique » (sgARN). De plus, il est le complémentaire de l'ARN génomique qui lui a servi de matrice, il est alors dit « anti-sens » (Sethna et al., 1989) (Figure 5.5). Après la production du sgARN anti-sens, le complexe de transcription/réplication se positionne sur l'extrémité 3' de ces molécules afin de produire le sgARN sens (Figure 5.6). Cet ARN est coiffé par l'action consécutive de plusieurs NSP. NSP13 hydrolyserait le phosphate γ présent sur le nucléotide qui forme l'extrémité 5' du nouvel ARN, avant l'ajout d'une guanosine monophosphatée par un acteur encore inconnu. Puis, NSP14 permettrait l'ajout d'un groupement méthyl sur la base guanosine de la coiffe, avant que NSP16, avec l'aide de son cofacteur NSP10, n'ajoute un groupement méthyl sur le premier ribose (position 2') de la séquence ARN (Aouadi et al., 2017). Cet ARN coiffé est également polyadénylé, probablement par l'activité 3' terminal adényl-transférase de NSP8 (Tvarogová et al., 2019). Le premier cadre de lecture du sgARN est alors traduit afin de produire les protéines virales non codées par orf1ab. Afin de produire toutes ces protéines, le complexe de transcription/réplication peut ignorer le premier TRS-b rencontré et poursuivre la transcription jusqu'à rencontrer un nouveau TRS-b au niveau duquel le choix de « sauter » ou de continuer est alors encore possible. En suivant ce processus, toute une série de sgARN sens est ainsi créée, chacun possédant la même extrémité 3' et la même extrémité 5' que le génome viral, mais dont le premier cadre de lecture (dans le sens 5' - 3') diffère en fonction des sgARN (Figure 5.6). Ces ARN sont donc structurellement polycistroniques (ils contiennent l'information génétique de plusieurs gènes), mais sont fonctionnellement monocistroniques (seul le premier cadre de lecture rencontré par les ribosomes est traduit), à l'exception du sgARN exprimant les protéines accessoires 3a, 3b et E [3c] (Liu et al., 1991; D. X. Liu and Inglis, 1992) et possiblement du sgARN permettant l'expression des gènes 5a et 5b (D.X. Liu and Inglis, 1992).

Lorsque le complexe de transcription/réplication ignore toutes les séquences TRS-b, alors la transcription complète du génome est réalisée, aboutissant à la création d'un ARN génomique anti-sens (Figure 5.7). Le complexe de transcription/réplication se fixe ensuite sur l'extrémité 3' de l'ARN génomique anti-sens et commence sa transcription. Cette transcription correspond à la synthèse d'un nouveau génome viral qui sera également coiffé en 5' et polyadénylé en 3'. Ce nouveau génome est alors disponible pour un nouveau cycle de transcription/réplication.

Le mécanisme permettant au complexe de transcription/réplication de « sauter » d'un TRS-b vers le TRS-L est le sujet de nombreuses recherches. Des interactions ARN/ARN s'opérant entre séquences fortement éloignées les unes des autres, ainsi que des interactions protéine/ARN sont supposées permettre le rapprochement du TRS-L avec les différents TRS-b. De ces rapprochements, un important encombrement stérique serait engendré et permettrait de stopper l'avancée du complexe de transcription/réplication juste après la transcription de la séquence TRS-b. La séquence complémentaire du TRS-b qui a été générée serait alors transférée sur la séquence TRS-L, où elle peut également s'hybrider, afin de pouvoir reprendre le processus de transcription (Sola et al., 2015). Cependant, la façon dont le complexe polymérasique choisit de « sauter » ou d'ignorer le TRS-b n'est toujours pas clair. Une autre hypothèse de mécanisme repose sur un des rôles de la protéine N. Dans le cas d'infection par IBV, la protéine N peut se localiser dans le noyau des cellules afin de permettre l'export de l'hélicase cellulaire DDX1 ainsi que son recrutement au niveau du complexe viral de transcription/réplication. L'utilisation de DDX1 à ce niveau permettrait d'éliminer l'encombrement stérique induit par les interactions ARN/ARN et ARN/protéine au niveau des TRS, pour que le complexe de transcription/réplication ne s'arrête pas au niveau des TRS et poursuive la transcription du génome (Wu et al., 2014).

Ce modèle de transcription/réplication discontinue du génome des coronavirus est celui qui a initialement été proposé en 1995 (Sawicki and Sawicki, 1995) et qui a par la suite été accepté et conforté par le reste de la communauté scientifique.

2.4.4 Assemblage/Bourgeonnement

L'assemblage consiste à réunir le génome viral avec l'ensemble des protéines de structure du virus et le bourgeonnement consiste à la formation de la particule virale finale (Figure 5, étapes 8, 9 et 10). Cependant ces étapes sont encore très mal connues chez les coronavirus (Masters, 2019). Dans le sous-genre *Embecovirus* des *Betacoronavirus* (appelé lignée A dans la publication), l'assemblage du génome viral et des protéines de structure fait intervenir des structures ARN particulières localisées au milieu de l'orflab dans la région codant pour NSP15 (chez le coronavirus hépatique de la souris). Des séquences similaires ont également été prédites dans les génomes des coronavirus de lapin et de cheval (deux *Betacoronavirus*) au sein de la région codant pour NSP3. Ces structures sont absentes de tous les autres types de coronavirus. Chez un *Alphacoronavirus* (coronavirus de la gastro-entérite porcine), une portion de la séquence 5'UTR a été identifiée comme permettant l'empaquetage, et au sein de cette région, des structures tige/boucle de l'ARN avec des séquences particulières ont été prédites comme pouvant être le signal d'empaquetage (Masters, 2019). Ainsi, il semble qu'il n'existe pas de signal d'empaquetage

universel commun à tous les coronavirus, ni même de signal commun au sein d'un même genre de coronavirus. Concernant les *Gammacoronavirus* et les *Deltacoronavirus*, aucune information n'est actuellement disponible à ce sujet, exception faite d'une étude sur IBV (Dalton et al., 2001) qui n'a pas réussi à identifier les régions nécessaires à l'empaquetage du génome viral.

L'interaction de la protéine M avec les protéines S, E et N assure le fait que tous les acteurs sont bien présents au moment du bourgeonnement pour former la particule virale. Le bourgeonnement s'effectue au niveau du compartiment intermédiaire entre le réticulum endoplasmique et l'appareil de Golgi (ERGIC – *Endoplasmic Reticulum Golgi Intermediate Compartment*). Cette étape fait intervenir la coopération des protéines E, M et N, car il a été montré que la production de particules virales dépourvues de génomes (VLP – *Virus-Like Particles*) était possible grâce à la co-expression des protéines E et M, potentialisée par la présence de la protéine N (Corse and Machamer, 2000; Ruch and Machamer, 2011; Liu et al., 2013). Tout en bourgeonnant, l'incorporation du génome viral est effectuée dans la particule virale.

Après le bourgeonnement, la particule virale se retrouve alors dans la lumière du ERGIC. Elle est ensuite transportée vers l'appareil de Golgi, puis vers la membrane plasmique, pour sortir de la cellule par exocytose (fusion des vésicules de transport [vésicule d'exocytose] avec la membrane plasmique) (Figure 5.8).

2.5 Méthodes de culture des coronavirus aviaires

2.5.1 Propagation sur culture cellulaire

Des méthodes de culture cellulaire ont été développées en utilisant des cellules primaires rénales ou de poumons afin de permettre la multiplication des souches historiques IBV-Mass41 et Beaudette (Cunningham and Spring, 1965; Lukert, 1966), mais beaucoup de souches requièrent une période d'adaptation avant de pouvoir se répliquer sur cellules. Un autre système a été développé pour la souche IBV-Beaudette, utilisant des cellules issues d'embryon de dinde, mais pour que cela puisse fonctionner, il a fallu d'abord effectuer quelques passages de la souche sur des cellules primaires de rein issues d'embryons de poulet (Coria and Peterson, 1971).

Des systèmes permettant la culture de souches d'IBV sur des lignées cellulaires immortelles ont également été développés. À plusieurs reprises, la souche IBV-Beaudette a été adaptée à la culture sur cellule Vero (cellule de reins de Singe Vert - *African green Monkey*) : une première fois à la suite de trois passages sur cerveau de souris et deux passages sur cellule Vero (Cunningham et al., 1972), et les autres fois à la suite de 10 ou 65 passages sur cellules Vero (Cunningham et al., 1972; Alonso-Caplen et al., 1984; Ng and Liu, 1998; Shen et al., 2003). Une autre étude a utilisé la lignée cellulaire CER (*Chicken Embryo Related cell line*) afin de cultiver la

souche IBV-Mass41, après seulement 5 passages d'adaptation (Ferreira et al., 2003). Les oiseaux n'étant pas des mammifères, il pourrait être imaginé que l'utilisation de cellules immortelles aviaires soit plus adaptée pour la culture d'IBV *in vitro*, afin d'éviter un possible changement génétique à la suite d'un saut d'espèce provoqué *in vitro*. Toutefois, la culture d'IBV sur cellules Vero a permis une meilleure compréhension des mécanismes moléculaires inhérents aux coronavirus aviaires et plus largement aux coronavirus. Mais bien que les lignées cellulaires immortelles soient plus simples à utiliser et à maintenir en culture, toutes les souches d'IBV ne se répliquent pas sur ces lignées, comme la souche IBV-QXD388. C'est pourquoi la culture de cellules primaires reste encore utilisée.

Dans le cas du TCoV, peu de souches ont été isolées à travers le monde. L'adaptation de la souche Minnesota sur la lignée immortelle HRT-18, issue d'un carcinome humain, est le seul système de culture cellulaire qui a été décrit à ce jour (Dea et al., 1991, 1990, 1989a, 1989b, 1989c), mais en dehors de cette équipe, ce système semble avoir été peu utilisé (Guy et al., 2000a).

Aucun système de culture cellulaire *in vitro* n'a été développé et testé pour le GFCoV.

2.5.2 Propagation sur explants

La propagation de l'IBV a également été testée sur des organes *ex vivo*, aussi appelés explants. Un premier système est la culture d'anneaux de trachée de poulet. Dans ce système, des poulets EOPS sont euthanasiés et les trachées sont collectées. Les anneaux cartilagineux, qui donnent sa structure à la trachée, sont séparés les uns des autres par coupure mécanique et sont ensuite incubés dans du milieu de culture à 37°C pendant environ 4 jours. La réplication du virus dans les cellules ciliées et sécrétrices de mucus peut être évaluée grâce à l'arrêt d'activité des cils vibratiles observables à l'aide d'un microscope. Cette méthode de culture sur organe est la plus utilisée mais des cultures utilisant des explants de rein, de proventricule et de bourse de Fabricius ont également été testées avec succès pour la culture de l'IBV, à la différence des explants d'intestins (Darbyshire et al., 1976; Darbyshire, 1978; Bhattacharjee and Jones, 1997).

À ce jour, il semblerait que la culture de TCoV et de GFCoV sur explant n'ait pas été testée. Mais des organoïdes intestinaux de dinde et de pintade pourraient être développés, à l'instar de ceux utilisés lors d'une étude sur le virus de la diarrhée épidémique porcine (Li et al., 2018).

2.5.3 Propagation sur œufs

Parce que les systèmes de culture cellulaire ne permettent pas de répliquer toutes les souches virales, et parce que la préparation d'explants est un processus délicat qui requiert l'utilisation d'animaux, les œufs embryonnés sont également utilisés afin d'isoler et d'amplifier les *Gammacoronavirus* aviaires. Cette méthode est préférentiellement réalisée avec des œufs issus

d'animaux exempts d'organismes pathogènes spécifiés (EOPS), mais elle peut également être réalisée sur des œufs issus d'élevages conventionnels avec un haut statut sanitaire et ce afin d'éviter la présence d'autres agents pathogènes qui seraient isolés en même temps que la souche de coronavirus d'intérêt.

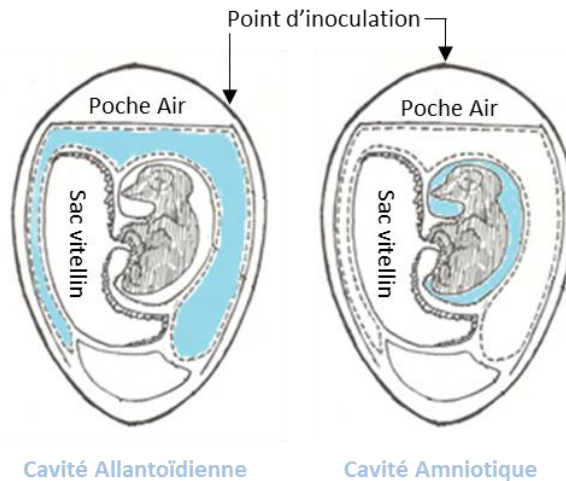


Figure 6: Description d'un œuf embryonné pour inoculation de coronavirus aviaires.

De récents travaux ont montré l'importance d'utiliser des œufs issus de la même espèce que celle chez qui le pathogène a été identifié (données non publiées). Ainsi, l'IBV peut être inoculé dans la cavité allantoïdienne d'œufs de poules embryonnés de 9 jours (Figure 6). La récolte du virus peut se faire dans un délai allant de 1 à 6 jours post inoculation. Durant cette période, les œufs sont incubés à 36,5°C. Passée la période d'incubation, le liquide allantoïdien est collecté et centrifugé afin de le clarifier, puis le surnageant est récupéré et utilisé comme stock viral (Keep et al., 2015). Concernant le TCoV ainsi que le GFCoV nouvellement isolé (données non publiées), ils peuvent être respectivement inoculés dans la cavité amniotique d'œufs de dindes et de pintades embryonnés de 19 jours (Figure 6). Suite à l'inoculation, les œufs sont incubés à 36,5°C pendant quatre jours. Dans les deux cas, la récolte du TCoV et du GFCoV se fait en collectant les intestins, comme précédemment décrit pour le TCoV (Guionie et al., 2013). Après homogénéisation des intestins dans un tampon salin et centrifugation, le surnageant est récupéré et utilisé comme stock viral.

3 Maladies et épidémiologie des *Gammacoronavirus* aviaires

3.1 La bronchite infectieuse causée par l'IBV

3.1.1 Maladies, espèces infectées et tissus cibles

L'IBV est l'agent étiologique de la bronchite infectieuse (BI), une maladie hautement contagieuse qui affecte les poulets et qui a été initialement décrite en 1931 par Schalk et Hawn aux USA (Fabricant, 1998). Aujourd'hui, après presque 90 ans de recherche, ce virus continue de provoquer de graves pertes économiques à l'industrie avicole (The World Bank, 2011).

La BI est une maladie infectieuse se caractérisant principalement par des troubles respiratoires, mais qui peut également provoquer des troubles rénaux et génitaux, du fait d'une réplication de l'IBV dans les tissus épithéliaux de différents organes (Doherty, 1967; Hofstad and Yoder, 1966). Indépendamment du type de souche, à tropisme respiratoire ou rénal, la maladie commence par se déclarer au travers de signes cliniques aspécifiques. Après l'infection, un inconfort généralisé apparaît rapidement, avec une perte d'appétit (anorexie) et une sensation de soif (hypodipsie) qui peuvent conduire à une perte de poids et à la déshydratation, suivies par de la toux, des râles, des écoulements nasaux ainsi que des bruits et des difficultés respiratoires. Certaines souches d'IBV peuvent provoquer des néphropathies et d'autres une baisse de la fertilité et une chute de ponte suite à une atteinte des oviductes (Jackwood and De Wit, 2013). La gravité des signes cliniques est variable selon l'âge des sujets. La maladie peut être indétectable chez les animaux de plus de six semaines d'âge, et les signes cliniques peuvent être sévères chez les animaux plus jeunes, en particulier autour de deux semaines d'âge (Animas et al., 1994a, 1994b; Khanh et al., 2018).

L'infection par IBV est l'une des infections les plus contagieuses actuellement décrite. Dans une population réceptive au virus et en absence de mesures de prévention, le taux de reproduction (R_0) calculé (correspondant au nombre moyen de sujets pouvant être infectés à partir d'un sujet infecté), est proche de 20 (De Wit et al., 1998). En dépit du fait que l'infection s'étende à tout l'élevage (morbidité de 100%), les taux de mortalité sont très variables, augmentés par des infections secondaires (le plus souvent bactériennes) et en présence de souches d'IBV néphropathogéniques (Jackwood and De Wit, 2013). Dès l'entrée du virus dans l'organisme du sujet commence une période d'incubation de courte durée, qui s'étend entre 18 et 36 heures en conditions expérimentales en fonction de la voie d'inoculation. Après son entrée dans l'organisme le virus commence à se répliquer et, selon le type de souche, peut se disséminer dans différents tissus de l'organisme. À ce jour, trois types de souches d'IBV représentent les souches majoritaires dans le monde : Massachusetts (plus communément appelé Mass, isolées dans les années 1930 aux

USA, (Beaudette and Hudson, 1937)), QX (isolées en 1996 en Chine, (YuDong et al., 1998)), 793b (isolées dans les années 1990 en Europe (Gough et al., 1992)).

Les souches de type Mass sont des souches à tropisme respiratoire avec un isolement possible du virus à partir de trachée infectée, des poumons ou des sacs aériens. Mais très tôt, la présence de virus infectieux de type Mass a été mise en évidence dans des organes tels les reins ou la bourse de Fabricius (Hofstad and Yoder, 1966). Avec des méthodes de biologie moléculaire, la présence d'ARN viral a pu être révélée dans les trachées, poumons, reins, intestins, oviductes et testicules, mais la détection d'antigènes viraux semble être restreinte aux organes des voies respiratoires (Benyeda et al., 2009). Cependant, des études ont également mis en évidence la présence d'antigènes viraux de souche Mass au niveau des épithéliums rénaux (Benyeda et al., 2010; Cheng et al., 2018), organe à partir duquel le virus a également pu être isolé (Fernando et al., 2017). D'autres études tendent à montrer que certaines souches de type Mass peuvent présenter un tropisme gonadique (testicules et ovaires) (Pereira et al., 2018) mais cette affirmation repose sur la détection de très faibles charges d'ARN viral qui sont considérées par les auteurs ayant mis au point la qRT-PCR comme étant inférieure à limite de détection génomique (Callison et al., 2006). Pour finir, à la suite d'une infection d'embryon de poulet avec une souche de type Mass, de l'ARN viral a été mis en évidence, grâce à l'hybridation *in situ*, dans tous les tissus épithéliaux de l'embryon qui ont été collectés (poumons, trachée, rein, oesophage, mésentère, bourse de Fabricius, sac aériens) (Lee et al., 2002).

Les souches de type QX ont été initialement décrites en Chine, où la première souche a été isolée à partir du proventricule de poulet (YuDong et al., 1998), organe situé en amont du gésier. Plus tard, ces souches ont de nouveau été détectées en Chine mais également en Europe. Dans les élevages, ou en conditions d'infections expérimentales, ces souches provoquent des bronchites ainsi que des néphropathies (Benyeda et al., 2010; Liu and Kong, 2004; Sigrist et al., 2012; Terregino et al., 2008), mais peuvent également provoquer des cystites (De Wit et al., 2011b).

Les souches de type 793b sont des souches à tropisme respiratoire, aussi bien en élevage que lors d'essais expérimentaux (Boroomand et al., 2011), mais initialement elles ont été détectées dans des élevages de poulets présentant des myopathies (Adzhar et al., 1995). De plus, la présence d'ARN viral a été détectée au niveau des reins, des testicules, des oviductes, des ovaires et des amygdales caecales (les plus larges organes lymphoïdes des intestins du poulet) (Boroomand et al., 2012). Par ailleurs, des souches 793b ont été isolées à partir de la trachée, des reins (lors de néphrite), et du cloaque (Mase et al., 2008). Ces souches ont également été montrées comme pouvant provoquer des cystites (De Wit et al., 2011b).

Les dénominations « respiratoire » et « rénale » des souches d'IBV reflètent la capacité des souches virales à induire une symptomatologie particulière, mais ne reflète pas la capacité des souches à infecter un tissu en particulier. En effet, une souche respiratoire peut être retrouvée dans d'autres organes que ceux du système respiratoire et une souche rénale n'est pas retrouvée uniquement au niveau du rein.

3.1.2 Classification des souches et épidémiologie

Devant le très grand nombre de souches isolées à travers le monde (Jackwood, 2012), et devant les multiples noms qui ont pu être donnés à une même souche (793b, aussi appelé CR88 ou 4/91), une règle de nomenclature a été proposée en 2001 afin d'harmoniser et de clarifier la dénomination des souches d'IBV (Cavanagh, 2001). Depuis 2016, sur proposition d'un groupe de travail européen (Ducatez and The European Union COST Action FA1207, 2016), la dénomination officielle de chaque nouvelle souche d'IBV suit la règle suivante :

genre_corona/AvCov/pays*/code_specimen**/année

(* = pays ou région d'un pays ; ** précédé de la lettre S (échantillon) ou I (isolat).

Par ailleurs, afin d'harmoniser la classification des génotypes, un système basé sur l'établissement d'un arbre phylogénétique utilisant la région S1 du gène S a été proposé (Figure 7) (Valastro et al., 2016).

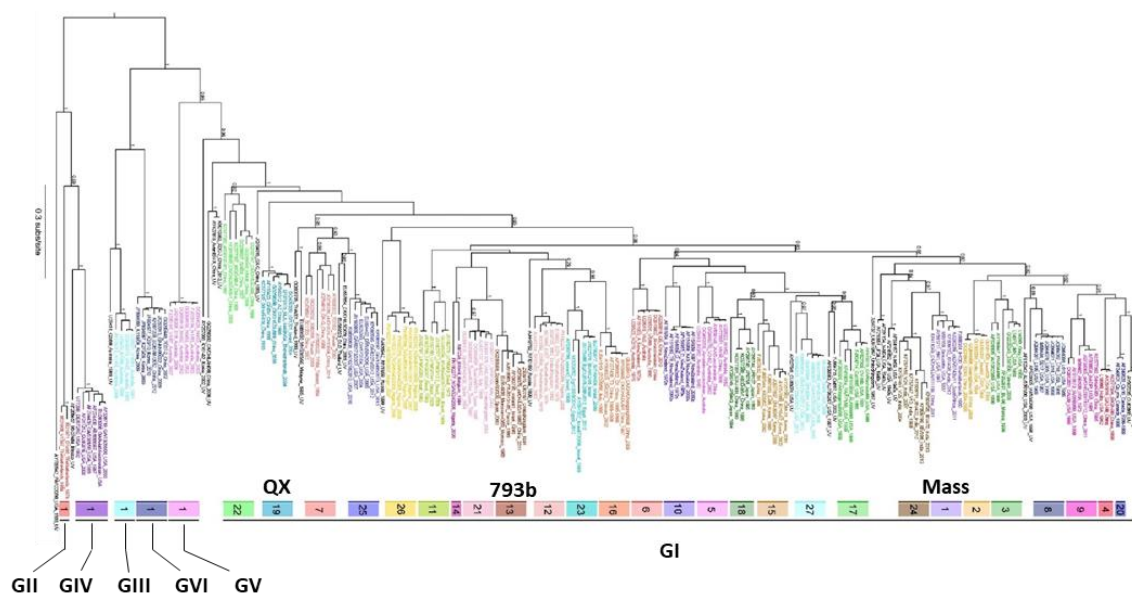


Figure 7 : Classification génotypique des souches IBV. Cet arbre phylogénétique utilise 199 séquences du gène S1 pour permettre la classification génotypique de toutes les souches IBV. Six génotypes, incluant différentes lignées, ont été identifiés. Image modifiée à partir de Valastro et al. 2016.

Parmi toutes les séquences de gène S1 de IBV, les auteurs ont retenu 199 séquences qui sont considérées comme étant les représentants de chaque génotype. Cette classification a permis d'établir six génotypes, notées GI à GVI, dans lesquels différentes lignées ont été identifiées. Ainsi, la comparaison de toutes séquences S1 nouvellement déterminées avec l'arbre phylogénétique permet de génotyper les nouvelles souches d'IBV. Les souches Mass, QX et 793b appartiennent respectivement aux lignées GI-1, GI-19 et GI-13. Récemment, une nouvelle souche IBV-D181 a été identifiée comme étant un nouveau sérotype IBV et le premier membre d'une seconde lignée du génotype GII (GII-2) (Molenaar et al., 2020). Les souches D1466 (GII-1) et D181 (GII-2) ont été identifiées comme provoquant des chutes de ponte (De Wit et al., 2020; Molenaar et al., 2020).

Depuis sa première description aux USA au début des années 1930, l'IBV s'est répandu à travers le monde. À ce jour le virus a été détecté dans tous les pays où il a été recherché sauf en Antarctique (De Wit et al., 2011a). Cependant toutes les régions du monde ne sont pas affectées par les mêmes souches d'IBV. Trois publications parues en 2011 (De Wit et al., 2011a), 2012 (Jackwood, 2012) et 2017 (Bande et al., 2017) font état de la répartition des souches d'IBV à travers le monde. Certaines de ces souches sont retrouvées dans la grande majorité des pays, c'est le cas des souches de type Mass, 793b et QX, mais d'autres souches dites locales sont uniquement retrouvées dans certains pays (Figure 8).

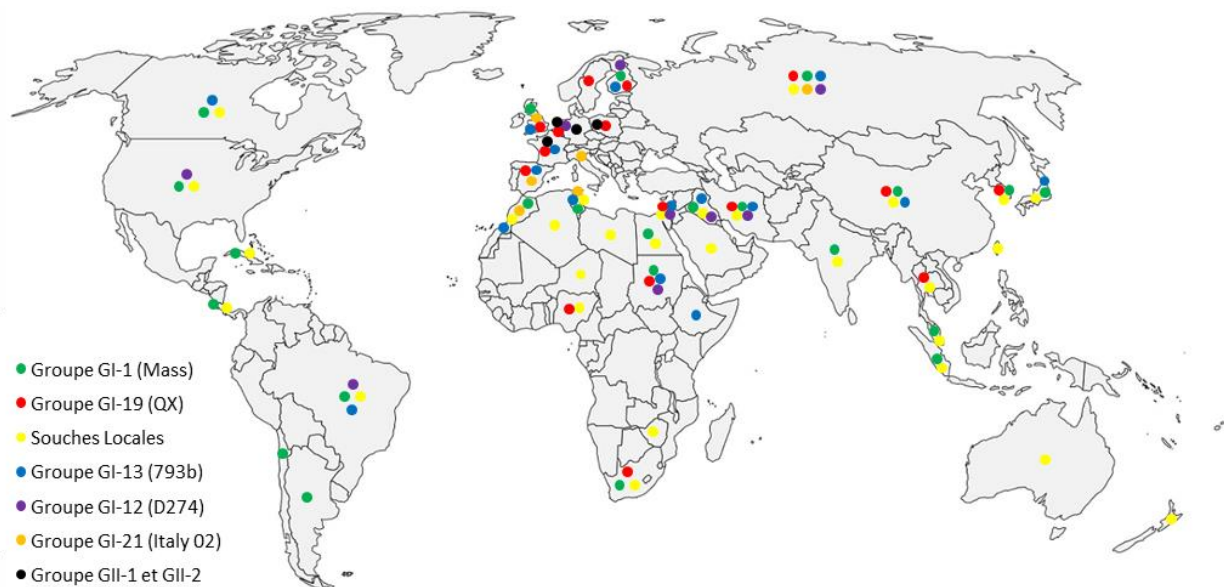


Figure 8 : Répartition géographique mondiale des différents génotypes IBV. Carte reprise et modifiée de Bande et al., 2017.

3.2 Coronavirose de la dinde causée par le TCoV

3.2.1 Maladie, espèce infectée et tissus cibles

Le complexe entérique de la dinde est une maladie hautement transmissible initialement décrite en 1951 sous la dénomination maladie de « Bluecomb » ou entérite transmissible de la dinde (Guy, 2017) et pour laquelle l'agent étiologique viral identifié est le TCoV (Panigrahy et al., 1973; Ritchie et al., 1973). Cette maladie d'origine multifactorielle, dont la survenue n'est pas inhérente à un seul pathogène mais est provoquée à la suite d'un contexte de co-infections, survient chez les dindes de tous âges mais préférentiellement avant sept semaines d'âge (Awe et al., 2015; Barnes et al., 2000). L'entérite de la dinde a pu être reproduite après inoculation du TCoV seul (Gomaa et al., 2009a, 2009b) ou en combinaison avec un *Escherichia coli* (Guy et al., 2000b; Pakpinyo et al., 2003), une bactérie entéropathogène, en permettant dans ces derniers cas d'obtenir des taux de mortalité très élevés (> 79 %). Ces résultats semblent concorder avec la réalité du terrain, où les contextes de co-infections sont fréquents (Moura-Alvarez et al., 2014, 2013) et où la sévérité de la maladie est souvent très élevée (Barnes et al., 2000).

Le complexe entérique de la dinde se caractérise par une importante réduction de croissance des sujets infectés du fait d'une dépression, d'une déshydratation, d'un manque d'appétit et d'importantes diarrhées. Les lésions macroscopiques sont une dilatation intestinale et un amincissement de la paroi intestinale induits par l'infection et la réplication du TCoV au niveau de l'épithélium intestinal (Barnes et al., 2000; Gomes et al., 2010; Guy et al., 1997), dans les parties moyennes et hautes des villosités intestinales (Gomes et al., 2010). Différentes études ont permis de montrer que les souches de TCoV détectées aux USA (TCOV-US) ont un tropisme préférentiel pour la partie basse de l'intestin, l'iléum et les caeca (Gomes et al., 2010; Guy et al., 2000b), et il semblerait que cela soit également vrai pour la souche Européenne (Fr-TCOV) (Brown et al., 2018). Par ailleurs, ces informations semblent coïncider avec la localisation d'un récepteur potentiel pour le TCoV (poly-N-acétyl-lactosamine) (Wickramasinghe et al., 2015).

Bien qu'à plusieurs reprises le TCoV ait été associé au complexe entérique de la dinde (Dea and Tijssen, 1988a, 1988b; Gomaa et al., 2009a; Jindal et al., 2014; Lin et al., 2002), cette maladie a également été décrite dans des élevages où le TCoV n'a pas été détecté (Jindal et al., 2010, 2009). Par ailleurs, dans une expérimentation impliquant la souche Européenne Fr-TCOV, la maladie n'a pas pu être reproduite (Brown et al., 2018).

Au cours d'une étude comparant l'infection de poulets et de dindes avec la souche TCoV VR911, plusieurs différences ont été observées. Après infection des poulets de un jour et de quatre semaines d'âge, aucun signe clinique n'a été observé au cours de l'expérimentation, mais le

génomique et des particules virales ont été retrouvés au niveau des intestins de poulets, le seul endroit où le virus a été cherché. Chez les dindes inoculées à un jour d'âge, des signes cliniques ont été observés, mais pas chez les sujets inoculés à quatre semaines d'âge (Ismail et al., 2003).

Au cours d'une autre étude utilisant des poulets inoculés à un jour d'âge avec la souche TCoV/Brazil/2006, l'ARN viral a été détecté dans les voies respiratoires et des antigènes viraux ont été retrouvés dans le sinus paranasal et la glande de Harder (ou glande de la membrane nictitante) (Gomes et al., 2010). En contradiction avec la première étude, ni le génome, ni les antigènes viraux n'ont été détectés dans les intestins des poulets. Cependant des souches différentes de TCoV ont été utilisées, ce qui pourrait expliquer cette différence.

3.2.2 Classification des souches et épidémiologie

Différentes souches de TCoV ont été décrites à travers le monde et isolées pour certaines, notamment sur les continents Américain et Européen (Cavanagh et al., 2001; Teixeira et al., 2007). À ce jour, deux lignées de TCoV peuvent être distinguées sur la base des génomes viraux excluant le gène S : les souches apparentées aux IBV d'origine américaine qui représentent la quasi-totalité des isolats, et les souches apparentées aux IBV d'origine européenne dont il n'existe pour l'heure qu'un seul isolat nommé Fr-TCoV 080385d (Brown et al., 2016). Les gènes S des deux lignées de TCoV (américaine et européenne) ont environ 60-65% d'identité nucléotidique en fonction des souches. Les gènes S des souches de TCoV européennes ont moins de 50% d'identité nucléotidique avec les souches d'IBV (Maurel et al., 2011).

Depuis l'isolement initial du TCoV européen, aucune donnée de répartition de ce virus n'est disponible. Une étude faisant état de TCoV circulant dans les élevages de dindes en Pologne ne permet pas de conclure à la détection de ce virus car les tests de détection du génome n'étaient pas spécifiques du TCoV mais des coronavirus aviaires et la caractérisation des génomes complets des coronavirus aviaires n'a pas été réalisée ne permettant pas de conclure à la circulation de TCoV en Pologne (Domanska-Blicharz et al., 2010).

Les paramètres de transmission de la souche TCoV européenne ont été évalués expérimentalement sans que le taux de reproduction (R_0) n'ait pu être calculé (Brown et al., 2018), mais cette étude a permis de calculer qu'en moyenne un sujet infecté pouvait en infecter un autre toutes les 2,5 heures. De plus, lors de cette expérimentation, les animaux infectés ont excrété le virus tout au long de l'expérience (une période de 6 semaines). Cette observation soulève la question de la persistance du TCoV chez les animaux infectés, question qui a déjà été soulevée avec des infections de poulets par l'IBV (Animas et al., 1994a). Cependant, cette longue période

d'excretion pourrait aussi être expliquée par des infections répétées des sujets avec les virus présents dans leur propre environnement.

3.3 Coronavirose de la pintade causée par le GFCoV

Très peu de données sont disponibles concernant le GFCoV car c'est un virus qui a été caractérisé récemment. En 2015, le premier génome complet de GFCoV a pu être déterminé à partir de pintades atteintes de la maladie X (Ducatez et al., 2015). Cette maladie, aussi appelée maladie foudroyante de la pintade, se manifeste par une prostration des sujets associée à des diarrhées verdâtres. Avec une morbidité de 100% et un taux de mortalité pouvant atteindre les 20% par jour, cette maladie se caractérise par des entérites aiguës pouvant être associées à des atteintes pancréatiques (telles que des hypertrophies) et des foies congestionnés voire hémorragiques (Berger, 2017). Depuis lors, le GFCoV est l'agent étiologique viral supposé de la maladie foudroyante de la pintade.

En 2017, sur des intestins provenant de pintades atteintes de la maladie foudroyante entérique, le séquençage à haut débit a permis de mettre en évidence un génome complet de GFCoV, mais également celui d'un picornavirus jusqu'alors inconnu (données non publiées). Après isolement du GFCoV, des essais expérimentaux ont été effectués sur des pintades conventionnelles de 26 jours d'âge. L'inoculation de l'isolat pur de GFCoV n'a pas permis de reproduire la maladie foudroyante, même si une légère prostration et de la diarrhée ont pu être observées chez certains sujets. En revanche, une entérite sévère associée à de la prostration a pu être observée après inoculation des pintades avec l'échantillon terrain de départ contenant les deux virus. Des études supplémentaires sont en cours afin d'étudier la capacité du picornavirus à reproduire la maladie foudroyante, mais de tels résultats semblent indiquer qu'à l'instar du complexe entérique de la dinde, il est possible que la maladie foudroyante de la pintade soit une maladie multifactorielle.

3.4 Les *Gammacoronavirus* dans d'autres espèces d'oiseaux

En plus de l'IBV, du TCoV et du GFCoV, qui infectent les poulets, les dindes et les pintades respectivement, d'autres *Gammacoronavirus* aviaires ont également été détectés et parfois isolés (Cavanagh, 2005) chez d'autres espèces telles que les cailles (Circella et al., 2007; Dubose et al., 1960; Torres et al., 2013), les faisans (Cavanagh et al., 2002; Gough et al., 1996; Spackman and Cameron, 1983), les pigeons (Felippe et al., 2010; Martini et al., 2018; Qian et al., 2006) et les perdrix (Cavanagh, 2005). Tous ces nouveaux coronavirus sont génétiquement liés à l'IBV et au TCoV.

3.5 La vaccination contre les *Gammacoronavirus* aviaires

Bien que la vaccination ne soit jamais efficace à 100% (Luginbuhl, 1952), c'est actuellement le meilleur moyen de limiter l'infection des élevages par l'IBV. Le premier vaccin anti-IBV a été réalisé en 1946 par van Roeckel et ses collègues (Jackwood and De Wit, 2013), basé sur la souche IBV-Mass41, qui était la seule souche circulante connue à ce moment là. Depuis, avec l'émergence de nouvelles souches d'IBV, pour lesquelles la vaccination contre Mass41 était inefficace du fait de changements antigéniques, de nombreux vaccins ont été développés pour une utilisation commerciale ou testés comme preuve de concept.

3.5.1 Les différents types de vaccins

Les vaccins anti-IBV se répartissent en trois catégories : les vaccins vivants atténués, les vaccins inactivés et les vaccins sous-unitaires (Jordan, 2017).

Les vaccins vivants atténués sont des souches virales adaptées à un système de culture (cellules, œufs) autre que l'hôte naturel et qui présentent un phénotype atténué après inoculation dans l'hôte naturel. Alors que le virus sauvage provoque la maladie, le vaccin vivant se réplique dans l'hôte à un niveau suffisamment faible pour ne pas provoquer de maladie, mais suffisant pour stimuler toutes les composantes du système immunitaire (humoral et cellulaire) du sujet vacciné.

Les vaccins inactivés sont des virus complets qui ont subi un traitement chimique ou thermique de façon à décomposer la particule virale, rendant toute infection impossible.

Les vaccins sous-unitaires peuvent être produits à partir d'une ou plusieurs protéines virales ou à partir de particules virales dépourvues de génome (*virus-like particles* (VLP)). Deux types d'approches peuvent être envisagées : i) l'utilisation de vecteurs d'expression permettant de produire des antigènes viraux afin de les administrer à un sujet dans le but de stimuler la composante humorale du système immunitaire (Eldemery et al., 2017) ; ii) l'administration de vecteurs d'expression (tel que des vecteurs adénovirus, des vaccins ADN réplcatifs ou non) dans des sujets pour permettre l'expression d'antigènes viraux *in cellulo* dans le but de stimuler la composante cellulaire de l'immunité adaptative, mais également de stimuler la composante humorale en cas d'excretion suffisante de la protéine dans le milieu extracellulaire (Johnson et al., 2003).

Les vaccins vivants semblent être majoritairement utilisés à travers le monde, alors que les vaccins inactivés semblent être le plus souvent utilisés en guise de rappel vaccinal (De Wit et al., 2011a; Jackwood and De Wit, 2013), notamment parce que l'utilisation de vaccin inactivé requiert l'inoculation individuel des sujets par voie intramusculaire, ajoutant à l'éleveur une charge de travail importante. L'immunité induite par l'administration de ces vaccins se révèle très efficace

pour prévenir les infections par des souches circulantes de même type (virus homologue au vaccin). En revanche, ils ne confèrent qu'une faible protection vis-à-vis d'une souche différente de celle utilisée lors de la vaccination (virus hétérologue au vaccin) (Lambrechts et al., 1993; Pensaert and Lambrechts, 1994; De Wit et al., 2011a), ce qui révèle l'importance de la surveillance et de la classification des souches virales circulantes, afin d'orienter au mieux le choix du vaccin.

3.5.2 Conséquences de certaines mauvaises pratiques vaccinales

Le développement des vaccins anti-IBV a été réalisé pour une utilisation dans des élevages de poulets, mais comme l'IBV a été détecté dans des élevages d'autres espèces aviaires, comme les cailles, certains éleveurs et vétérinaires ont également choisi de vacciner ces élevages avec des vaccins atténués dirigés contre l'IBV (Torres et al., 2017, 2016). Or, ces oiseaux peuvent également être infectés par des *Deltacoronavirus* (Torres et al., 2016) ou par d'autres *Gammacoronavirus*. La vaccination contre l'IBV pourrait alors engendrer la co-infection de ces oiseaux par un *Gamma-* et un *Deltacoronavirus*, ou par deux *Gammacoronavirus* augmentant les risques d'un retour à la virulence des souches vaccinales et l'émergence de nouvelles souches virales par recombinaison génétique entre les génomes. À ce jour, aucune recombinaison n'a été décrite entre des membres des *Gammacoronavirus* et des *Deltacoronavirus*.

Par ailleurs, en l'absence de réglementation uniformisée, de multiples protocoles de vaccination sont utilisés (Legnardi et al., 2019; Sadri et al., 2019), comme l'utilisation simultanée de plusieurs vaccins vivants atténués, pouvant ainsi favoriser l'émergence de nouvelles souches virales par recombinaisons génétiques entre les différentes souches inoculées (Hopkins and Yoder, 1986; Nix et al., 2000). Pour cette raison, ce type de protocoles vaccinaux est interdit en France.

3.6 Les méthodes de détection des coronavirus aviaires

Les premières détections d'infection par l'IBV se basaient sur la mise en évidence de particules virales infectieuses après inoculations successives d'échantillons sur des œufs de poules embryonnés jusqu'à l'obtention de lésions embryonnaires, ce qui selon le type d'échantillon ou de souche pouvait être plus ou moins long, notamment du fait que les souches pouvaient nécessiter quelques passages d'adaptation sur œufs de poule avant de pouvoir provoquer des lésions (Beaudette and Hudson, 1937; Cunningham and Stuart, 1947).

3.6.1 Méthodes de détection immunologique *in vitro*

Les méthodes immunologiques utilisent des anticorps ou des antigènes afin de mettre en évidence la présence d'antigènes ou d'anticorps spécifiques d'un agent infectieux.

En 1951, en reprenant des travaux de ses collègues, Fabricant a mis au point une méthode de détection d'anticorps anti-IBV nommé séroneutralisation (Fabricant, 1951). Pour cela, un stock d'IBV-Beaudette est inoculé en présence et en absence du sérum à tester. Si le sérum à tester contient des anticorps capables de neutraliser la souche IBV-Beaudette, alors le virus ne se répliquera pas après inoculation sur œufs embryonnés. Cette approche fonctionnelle présente le double inconvénient d'être longue à mettre en place et de n'être possible qu'en présence d'anticorps capables de reconnaître le sérotype viral. Sur le même principe, il est possible d'utiliser des séra connus pour neutraliser une ou des souches d'IBV afin de mettre en évidence la présence de particules virales dans l'échantillon à tester. Cette approche se nomme vironeutralisation. Elle présente l'avantage de pouvoir déterminer le sérotype de la souche d'IBV testée, mais nécessite de posséder une sérothèque contenant un grand nombre de séra afin de correctement caractériser la souche virale.

La technique de dosage d'immunoabsorption par enzyme liée, dite méthode ELISA (*Enzyme-Linked ImmunoSorbent Assay*) permet à partir d'une matrice liquide (sérum, écouvillon) de révéler la présence de l'agent pathogène viral ou d'anticorps qui lui sont spécifiques. Du fait de la multitude de sérotypes d'IBV, de nombreux tests ELISA ont été mis au point pour la détection de ce virus, mais peu ont été développés pour la recherche du TCoV et aucun n'existe pour le GFCoV.

Des approches d'immunohistochimie et d'immunofluorescence ont également été proposées (Jonsson and Engstrom, 1986; Lukert, 1966) en se basant sur la détection d'antigènes viraux dans des portions de tissus fixés.

Cependant, le manque de sensibilité des tests immunologiques a favorisé le développement d'outils moléculaires pour la détection de l'agent pathogène viral.

3.6.2 Méthodes de détection moléculaire

De nos jours, le diagnostic de maladies infectieuses repose sur la détection d'acides nucléiques viraux, qui à la différence des tests sérologiques, permet de mettre en évidence la présence du génome de l'agent pathogène recherché. Cette détection est basée sur l'amplification spécifique de portion de génome grâce à la réaction en chaîne d'une polymérase (PCR). Pour les virus à ARN, une transcription inverse du génome est d'abord nécessaire afin de produire un ADNc qui pourra ensuite être amplifié par PCR, il s'agit alors de RT-PCR (*Reverse Transcription - PCR*). Lorsque la RT-PCR est combinée à l'utilisation de marqueur fluorescent, comme des sondes fluorescentes capables de s'hybrider sur des séquences nucléotidiques spécifiques ou des molécules capables de s'insérer entre les bases de la double hélice ADN, il s'agit alors de RT-

PCR temps réel (*real time (rt)* RT-PCR). Cette méthode permet de déterminer en temps réel la présence d'acides nucléiques viraux dans un échantillon, et peut également permettre de quantifier cette présence si la méthode est couplée à l'utilisation de gamme d'étalonnage dont les concentrations en cibles ARN ou ADN sont connues. En 2006, une pan-rt RT-PCR quantitative (qRT-PCR) a été développée (Callison et al., 2006) afin de détecter l'ensemble des coronavirus aviaires actuellement connus : IBV et TCoV. Cette méthode permet également de détecter la présence de GFCoV, dont la caractérisation du génome est postérieure au développement de la qRT-PCR. Cette qRT-PCR est actuellement utilisée en routine par de nombreux laboratoires.

En revanche, en dépit du fait que certaines PCR permettent de mettre en évidence de multiples souches virales, les PCR requièrent l'usage d'amorces spécifiques. Cette spécificité empêche la détection de souches dont les séquences cibles du génome viral sont suffisamment éloignées pour empêcher l'hybridation des amorces sur le génome viral.

3.6.3 Méthode de détection par Séquençage Haut-Débit

Les séquenceurs de deuxième génération, aussi appelés NGS (*Next Generation Sequencing* en anglais) ou séquençage haut débit, permettent de déterminer rapidement le génome d'un organisme. En permettant de séquencer sans à priori des fragments de séquence nucléique, l'utilisation du NGS dans un objectif de détection permet de pallier le principal défaut de la RT-PCR, à savoir l'utilisation d'amorce spécifique s'hybridant sur des portions connues de génome.

Avec cette méthode, les génomes sont fragmentés par digestion enzymatique ou par sonication. Des adaptateurs (de courtes séquences ADN connues) sont ensuite liés aux extrémités des fragments de génome. Par la suite, les fragments peuvent être séquencés grâce à l'utilisation d'amorces qui s'hybrident sur les adaptateurs. Le séquençage fait intervenir une polymérase qui utilise un seul des deux brins du fragment de génome afin de générer le brin complémentaire et ainsi déterminer la séquence du fragment. La méthode de détermination des nucléotides ajoutés varie selon la technologie utilisée (détection des variations de pH ou des variations de luminosité). Les séquences déterminées à partir des courts fragments de génomes sont appelées des « lectures » (*reads* en anglais) (Figure 9).

Ces lectures sont traitées bioinformatiquement dans le but de recréer le génome séquencé. Ceci peut se faire sans à priori, c'est-à-dire en utilisant les lectures dans des processus d'assemblage *de novo* de façon à les organiser les unes par rapports aux autres via la recherche de zones chevauchantes, ou bien en organisant les lectures les unes par rapports aux autres via l'utilisation d'un génome de référence, sur lequel les lectures vont être alignées.

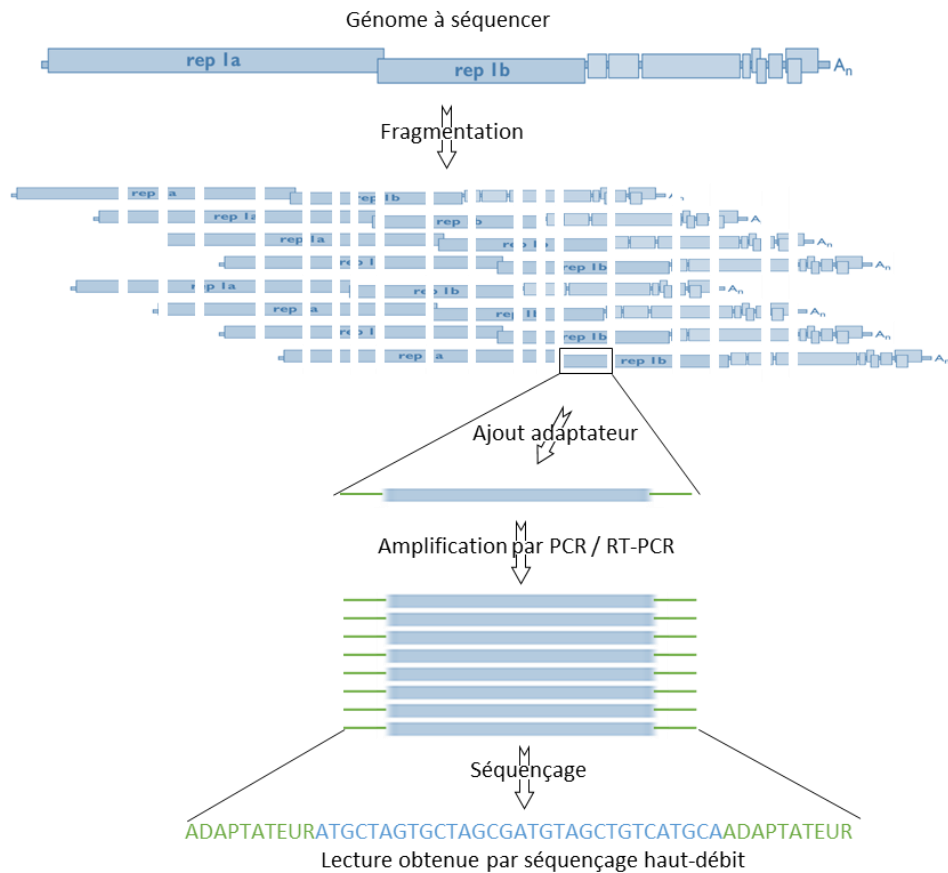


Figure 9: Processus de détermination de séquences par séquençage haut-débit.

Dans un objectif de détection d'agent pathogène, il n'est pas nécessaire de recréer bioinformatiquement le génome viral dans son intégralité. Les lectures peuvent être utilisées pour rechercher une correspondance dans les bases de données afin de déterminer à quel organisme appartient la lecture et ainsi déterminer le contenu de l'échantillon. Pour cela, différentes bases de données publiques contenant uniquement des génomes de virus ont été créées afin de faciliter et d'orienter la détection : ViPR (Pickett et al., 2012), Virosaurus (<https://viralzone.expasy.org/8676>) et RVDB (Goodacre et al., 2018).

La base de donnée ViPR regroupe des séquences de virus pathogènes, mais malgré le fait qu'elle ait été créée en 2012, cette base est toujours loin d'être complète. Virosaurus est une base de données plus récente qui a également été créée dans un but de détection, mais contrairement à ViPR, elle regroupe l'ensemble des séquences de virus de vertébrés actuellement connus, ce qui la rend plus complète. RVDB est une base de données plus imposante que Virosaurus. Elle a été pensée pour la découverte de nouveaux génomes viraux et regroupe non seulement des séquences virales, mais également des séquences de génome non viraux ayant une origine virale (comme des séquences endogènes rétrovirales ou encore des rétrotransposons).

4 Evolution des génomes viraux

4.1 Les théories de l'évolution

Le concept d'évolution est apparu au cours du XIXe siècle, au travers de deux conceptions différentes proposées par Jean-Baptiste Lamarck (1809) et Charles Darwin (1859). Pour résumer, la théorie de Lamarck repose sur la génération spontanée d'une forme de vie « simple » (individu) qui se complexifie à mesure des changements qui s'opèrent dans son environnement. Cette théorie raisonne sur l'adaptation d'un individu face aux changements de son environnement. Pour Darwin, les différents individus d'une même population présentent quelques différences les uns par rapport aux autres (c'est la diversité). Cette théorie raisonne sur l'adaptation d'une espèce (un ensemble d'individu d'une même population et leur descendance) face aux changements de son environnement, via une sélection d'individus capables de se reproduire malgré les modifications de leur environnement. Darwin introduit ici la notion de pression de sélection, une force capable de provoquer la survie (sélection) de certains individus (et donc la mort d'autres) afin conserver les caractéristiques qui permettent aux individus de survivre dans la population.

4.2 Création et transmission de la diversité

En 1978, Esteban Domingo révèle qu'un isolat pur de bactériophage Q β (génome ARN) est en réalité génétiquement hétérogène, indiquant qu'une minorité de la population de génome possède des différences génétiques vis-à-vis de la majorité (Domingo et al., 1978). Cette découverte, a permis d'établir le concept de « quasi-espèce virale », et plus tard renommée concept de « population virale ». Ce concept fait résonance à la théorie de Darwin sur l'évolution, en ce sens où la population d'un seul isolat viral est composée de génomes (des individus) qui ne sont pas identiques aux autres génomes de l'isolat mais dont les petites différences permettent d'adapter la population virale (l'espèce virale) aux changements de son environnement. Dans le génome de coronavirus, ces mutations peuvent survenir à la suite de différents phénomènes :

- des réactions chimiques spontanées : alkylation des bases (ajout de groupement méthyl, ethyl ou propyl sur les bases), déamination des cytosines en uracile, déamination des méthyl-cytosines en thymine.
- des erreurs induites par la polymérase virale : lors de la réplication du génome ARN viral, l'extension du nouveau brin ARN est catalysé par l'ARN polymérase virale. Ces enzymes ou complexes enzymatiques sont sujets à erreur (comme toute les polymérases cellulaires ou virales). Les ARN polymérases sont moins fidèles que les ADN polymérases et produisent donc plus d'erreurs lors du processus de réplication des

génomés. Cette différence de fidélité est accentuée par le fait que la grande majorité des polymérase virales ne sont pas associées à des systèmes de correction d'erreurs à la différence des ADN polymérase cellulaires. Le seul système de correction actuellement décrit pour les virus à ARN, repose sur la présence d'une activité 3'-5' exonucléasique contenue dans une sous-unité du complexe de transcription/réplication du génome des *Coronaviridae* (Snijder et al., 2016). La présence de ce système permet de réduire la survenue d'erreur, sans l'annuler.

- les recombinaisons : il s'agit de l'échange de portion de génome entre deux virus infectant une même cellule. Ces échanges peuvent se faire en échangeant un même segment de génome (recombinaison homologue) ou bien en échangeant des séquences de natures différentes (recombinaison hétérologue).

Cette capacité à générer des erreurs, associée à la génération de plusieurs millions de copies de génome viraux par cycle, contribue à l'augmentation de la diversité génétique d'une population virale. En utilisant la génétique inverse afin d'augmenter ou de réduire la fidélité de l'ARN polymérase virale, permettant ainsi de réduire ou d'augmenter la complexité de la diversité génétique, une diminution de la capacité d'adaptation du virus a été constatée face à un changement environnemental donné (Vignuzzi et al., 2008). Un niveau optimal de diversité génétique est donc supposé nécessaire pour permettre une bonne réplique du virus et une bonne adaptation de la population virale face à de nouveaux environnements.

Afin que les variations génétiques soient maintenues dans une population, il est nécessaire qu'elles soient transmises aux nouveaux individus de la population. Dans le cas d'une infection d'un être vivant par un virus, le maintien de variants génétiques passe également par la transmission de l'agent viral à un autre être vivant. Après l'entrée du virus dans l'organisme, commence une période de latence² plus ou moins longue pendant laquelle l'infection virale s'étend aux tissus sensibles de l'organisme et s'y réplique, afin d'augmenter la quantité de matériel viral au sein de l'organisme infecté et de générer de la diversité virale. À la fin de cette période, le virus est excrété dans l'environnement en quantité suffisamment importante pour infecter un autre organisme et à chaque nouvelle infection le processus recommence. Mais lors de la transmission, la totalité des virus produit par l'organisme infecté n'est pas transmise à un nouvel organisme, si bien que seule une fraction de la diversité génétique, correspondant à la majorité de la population virale au moment de la transmission, est transmise. Ceci est l'effet du goulot d'étranglement.

² Délai entre l'infection d'un sujet et le moment où il devient contagieux, à la différence de l'incubation qui se termine par l'apparition des signes cliniques

4.3 Le système immunitaire adaptatif comme pression de sélection

À la suite d'une primo-infection, lorsqu'un organisme rencontre un agent infectieux pour la première fois, un délai d'activation du système immunitaire adaptatif est nécessaire afin qu'il soit opérationnel. À la suite de cette activation, des anticorps et des cellules immunitaires spécifiques de l'agent infectieux sont produits. Lors d'une infection ultérieure avec le même agent pathogène, le délai d'activation de la réponse immunitaire adaptative est réduit puisque les effecteurs de cette réponse sont déjà présents dans l'organisme permettant ainsi une réponse ciblée plus rapide que lors de la primo-infection. La vaccination permet de mimer une primo-infection de sorte à stimuler une première fois le système immunitaire adaptatif dans le but de réduire son délai d'activation lors d'infections futures et de réduire la durée et la sévérité de l'infection. Deux composantes forment le système immunitaire adaptatif : la composante humorale et la composante cellulaire.

La composante humorale permet de développer des anticorps dirigés spécifiquement contre des épitopes, des régions spécifiques de protéines (virales quand il s'agit de virus). Ces épitopes peuvent être de deux types : des épitopes linéaires, formés par une succession définie d'acides aminés, ou bien des épitopes conformationnels, formés par des acides aminés pouvant être distants mais qui deviennent proches lors de la mise en place des structures secondaires, tertiaires et/ou quaternaires des protéines. Lorsque les anticorps sont dirigés contre des protéines de surfaces des particules virales, alors ils se fixent sur celles-ci. Si l'épitope de la protéine recouvert par l'anticorps était le déterminant antigénique responsable de l'entrée du virus dans la cellule, alors la protéine virale ne pourra pas remplir sa fonction et le virus sera incapable de pénétrer dans la cellule. Ainsi, en présence d'une quantité suffisante d'anticorps capables de neutraliser un virus, toutes les particules virales pourront être neutralisées et l'infection sera stoppée.

Mais dans le cas d'une population virale, dont les génomes peuvent varier, il peut arriver que les variations génétiques de la population entraînent l'apparition de protéine de surface dont la séquence a été modifiée, entraînant possiblement un changement de l'épitope qui n'est alors plus reconnu par les anticorps neutralisants (Figure 10). Ainsi, lors d'une infection virale chez un sujet possédant déjà des anticorps contre l'agent infectieux, les particules virales ne possédant pas cette modification seront neutralisées, et une sélection des particules virales avec des protéines de surface modifiées sera effectuée, car seules ces particules seront capables d'infection et donc de réplication. La population virale résultante aura donc évolué à la suite d'une pression de sélection induite par les anticorps.

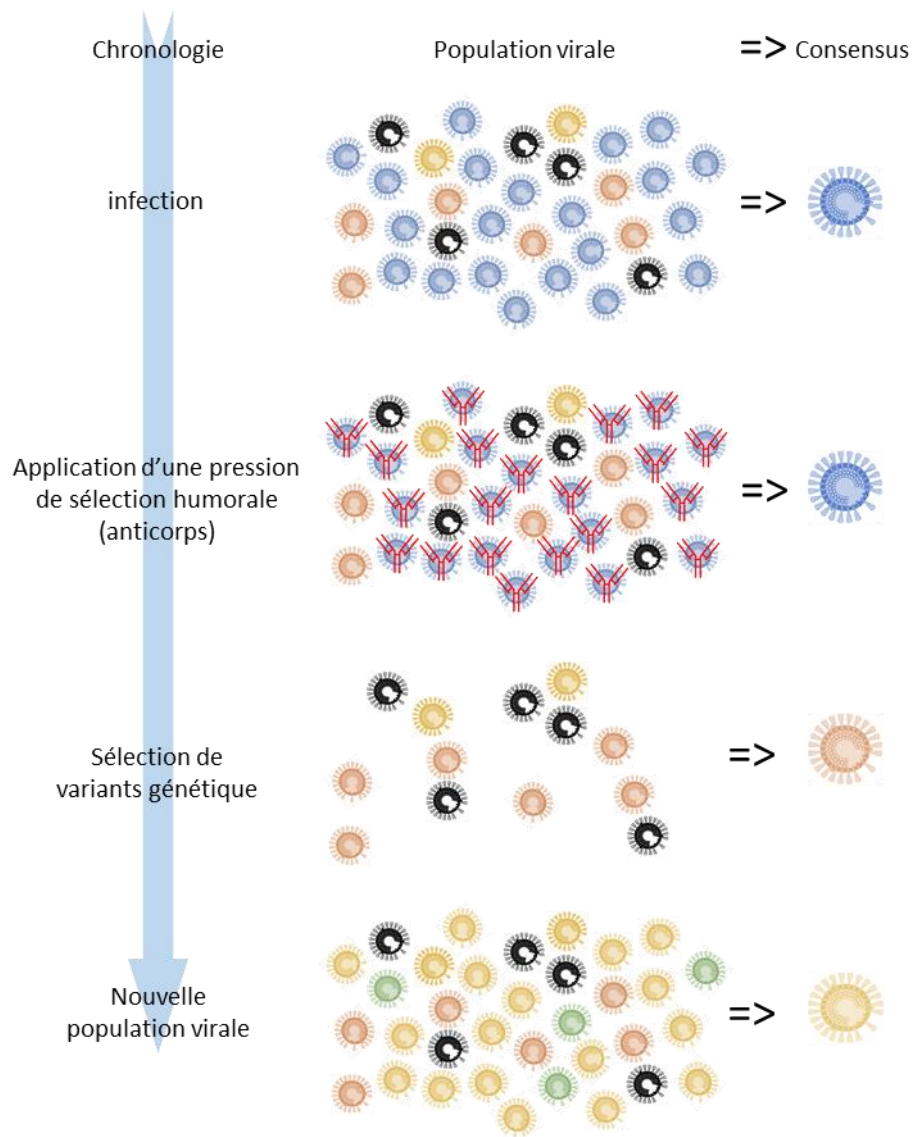


Figure 10 : Illustration d'une sélection de population virale par l'intermédiaire des anticorps neutralisants.

Le même phénomène de sélection peut se produire avec la composante cellulaire de l'immunité adaptative. Cette composante analyse le contenu protéique des cellules via la présentation de peptides au sein des molécules du complexe majeur d'histocompatibilité de type I (CMH-I), exposées à la surface cellulaire. Ces peptides sont des fragments de protéines générés par la cellule lors de la dégradation des protéines cellulaires. Ainsi, lors d'une infection virale, les protéines virales présentes dans la cellule et dans les membranes cellulaires seront également dégradées et présentées dans les molécules du CMH-I. Cependant chaque molécule du CMH-I ne peut pas présenter toutes les séquences peptidiques possibles, ce qui fait que l'utilisation de multiples molécules du CMH-I permet de présenter un large repertoire de peptides, mais pas la totalité. Ainsi, lorsqu'un virus sauvage présente des peptides capables d'être présentés dans les molécules du CMH-I, la présence du virus sera alors détectée par les cellules immunitaires et la

cellule infectée pourra être détruite, permettant de stopper l'infection. Mais si dans la population virale, des variants génétiques à partir desquels les peptides ne seront plus présentés dans les molécules du CMH-I sont retrouvés, alors ces variants génétiques ne seront pas reconnus par les cellules du système immunitaire et l'infection se poursuivra. Une sélection de génomes viraux capables de produire les protéines variantes est alors effectuée.

4.4 Analyse de l'évolution des génomes

Initialement, la constatation de l'évolution des génomes viraux se faisait sur la base des séquences consensus virales, déterminées par la méthode de séquençage utilisant des terminateurs de chaînes (méthode Sanger). Mais du fait du coût important du séquençage, cette méthode était fréquemment réalisée sur une portion du génome viral. De plus en plus, ce type d'analyses utilise le séquençage haut débit, qui permet idéalement d'obtenir la séquence d'un génome complet pour un prix relativement bas. Avec cette technologie, la séquence consensus du virus peut être déterminée et comparée aux séquences existantes via l'utilisation des bases de données. Mais comme chaque lecture provenant du NGS correspond à un fragment de molécules génomiques, lorsqu'un variant génétique est présent dans une population virale, ce variant va se retrouver dans une ou plusieurs des lectures du NGS. En théorie, cette technologie permet donc de détecter tous les variants génétiques d'une population virale, via l'utilisation d'outils bioinformatiques dédiés.

D'une façon générale, les outils de recherche de variants génétiques ne permettent pas d'utiliser directement les données issues de séquençage haut débit, mais utilisent des fichiers d'alignement de lectures sur des génomes de référence, ce qui nécessite des compétences en bioinformatique. En 2015, un outil bioinformatique, nommé ViVan pipeline (Isakov et al., 2015), ne nécessitant pas de compétences particulières en bioinformatique, et donc est utilisable par tous les scientifiques, a été publié, mais cet outil n'est plus disponible.

5 Problématique de la thèse

Alors que de nombreux vaccins existent pour limiter les infections causées par IBV, une grande diversité antigénique existe pour ce virus. Malgré des décennies de vaccination intensive des élevages de poulets dans le monde, les infections à IBV sont encore très fréquentes et de nouvelles souches virales continuent d'être caractérisées dans différents pays. En contraste, aucun vaccin n'existe pour limiter les infections TCoV et seulement deux grands types de souches ont été recensés : les souches TCoV-US et la souche Fr-TCoV.

L'apparition des premiers vaccins anti-IBV a permis de protéger les individus vaccinés face au seul sérotype identifié circulant à l'époque, le sérotype Mass. Plus tard, une nouvelle souche virale a été isolée à partir d'animaux vaccinés développant des signes cliniques de BI. Cette nouvelle souche, Connecticut, s'est révélée être un nouveau sérotype, à partir duquel un nouveau vaccin a été développé. Ce phénomène s'est répété plusieurs fois jusqu'à aujourd'hui. L'émergence des nouvelles souches d'IBV est le résultat de l'évolution des génomes viraux, dont les modifications sont générées au cours de la réplication du virus dans les sujets infectés. En accord avec la théorie de la Reine Rouge³, qui évoque le besoin d'évolution permanente d'une espèce afin d'adapter sa survie en fonction de celle des autres espèces dans le même environnement, il a été supposé que l'usage de la vaccination aurait entraîné une pression de sélection capable de provoquer l'émergence de nouvelles souches virales, qui ont acquis la capacité de se répliquer en présence d'une immunité anti-IBV appartenant à un autre sérotype. Cependant, il ne faut pas rejeter l'hypothèse selon laquelle les vaccins n'auraient pas favorisé l'émergence des nouvelles souches, mais auraient servi à révéler la présence de souches circulantes qui étaient encore inconnues. Les nouveaux sérotypes viraux auraient pu apparaître et coexister avec les « anciens », sans que l'Homme n'ait eu conscience de cette diversité, jusqu'à l'introduction de vaccins qui a permis de révéler des différences de protection des animaux vaccinés vis-à-vis de différentes souches d'IBV. Cela peut d'ailleurs être vrai encore aujourd'hui, avec de nouvelles souches d'IBV qui sont souvent caractérisées à la suite d'échecs de vaccination.

Devant ces observations et questionnements, ce projet de thèse a été déposé afin d'étudier l'évolution des *Gammacoronavirus* aviaires TCoV et IBV au cours de multiples passages des virus sur dindes ou poulets, en absence ou en présence d'une immunité vaccinale pré-existante. Les variants génétiques qui composent les populations virales des deux virus ont été caractérisés par séquençage haut-débit afin d'évaluer le dynamisme des populations virales de TCoV et d'IBV.

³ Cette théorie énoncée par Leigh van Valen fait référence à la Reine rouge du roman de Lewis Carroll de 1871, Through the Looking Glass : "Now, here, you see, it takes all the running you can do, to keep in the same place. If you want to get somewhere else, you must run at least twice as fast as that!" - *The Red Queen*

Deuxième partie
-
Travaux Personnels

6 Développement d'un outil de visualisation des variants génétiques viraux

La première étape de ce projet de thèse reposait sur la capacité de détection et de visualisation simple des variants génétiques viraux issus de données de NGS et s'est donc focalisé sur le développement d'un outil permettant d'atteindre cet objectif.

Les outils de recherche de variants (*variant caller*) utilisent des données de séquençage qui ont déjà subi un traitement bioinformatique par des outils d'alignement des données de séquençage sur des génomes de références. De plus, les résultats issus des « *variant callers* » sont des fichiers texte difficilement analysables sans connaissance préalable du format de fichier. Dans ce contexte, il est apparu que les outils disponibles n'étaient pas facilement utilisables par des non-bioinformaticiens. Une publication parue en 2015 faisait état d'un outil bioinformatique nommé ViVan pipeline (Isakov et al., 2015) qui présentait l'avantage d'une utilisation simplifiée en transférant des données de NGS, ainsi qu'un génome de référence, à un serveur de calcul via un navigateur internet. Cependant, cet outil n'était plus utilisable. Il a donc été décidé de procéder à la création d'un outil simple d'utilisation permettant de réaliser toutes les étapes nécessaires au traitement bioinformatique et d'obtenir une vue simplifiée des résultats sous forme de graphiques. Dans cette optique, une succession d'étapes bioinformatiques (workflow) a été mise en place afin de procéder au traitement des données de NGS. Pour cela, Snakemake, un gestionnaire de workflow à la syntaxe basée sur Python, a été utilisé afin de pouvoir organiser et répéter toutes les étapes du processus de traitement des données. L'outil final est utilisable grâce à une instance GALAXY en ligne via l'url suivant : <https://viralvariant.anses.fr>. Cet outil a été nommé Viral Variant Visualiser (VVV) et il peut être utilisé de deux manières (Figure 11).

La première utilisation consiste à fournir à VVV les données de séquençage uniquement (Figure 11, tracé vert et bleu). L'outil procédera alors à l'identification du virus le plus présent dans l'échantillon et recréera la séquence majoritaire du génome viral (séquence consensus). Cette séquence consensus est ensuite utilisée pour aligner les lectures NGS et procéder ainsi à la recherche de variations génétiques qui pourraient composer la population virale associée à ce génome.

La seconde utilisation consiste à fournir à VVV les données de séquençage ainsi que la séquence d'un génome viral (Figure 11, tracé vert et jaune). La séquence fournie sera utilisée pour aligner les lectures NGS et procéder directement à la recherche de variations génétiques qui pourraient composer la population virale associée au génome fourni. Cette seconde utilisation présente en plus l'avantage de pouvoir comparer n'importe quelle séquence de génome avec les données NGS fournies.

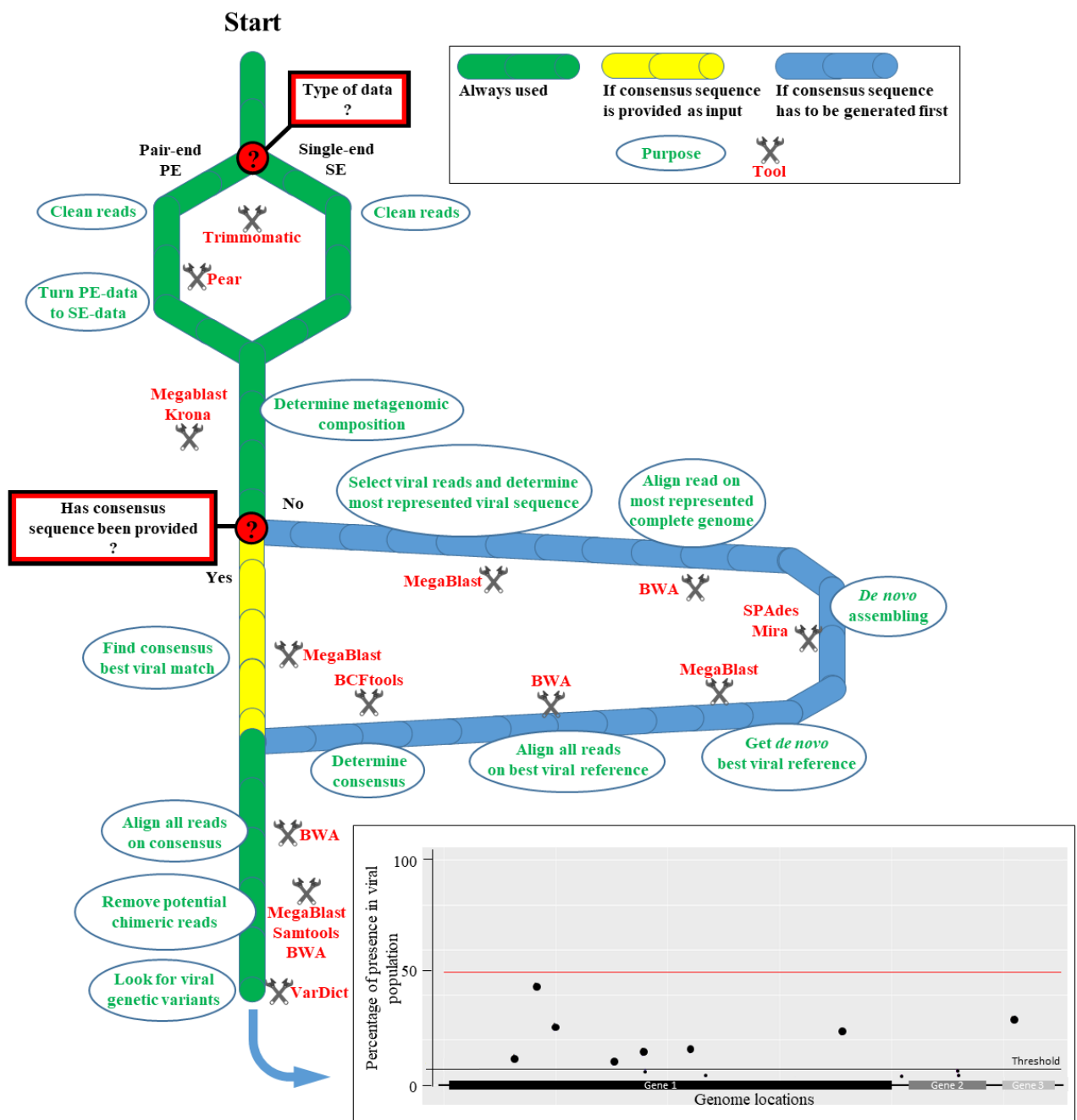


Figure 11: Idéogramme décrivant le processus de traitement des données par l’outil Viral Variant Visualiser.

La visualisation des variants génétiques s’effectue sur un graphique qui présente en abscisse les différentes positions du génome viral et en ordonnée le pourcentage de présence des variants génétiques dans la population virale à chaque position. Tout variant génétique présent au-dessus d’un seuil spécifié par l’utilisateur (qui est un pourcentage minimal de présence dans la population virale, seuil fixé par défaut à 7%) apparaîtra sous la forme d’un point dont la hauteur définira son pourcentage de présence dans la population virale.

Pour tester VVV, ce seuil de positivité des variants génétique de 7% a été utilisé, afin de limiter la caractérisation d'erreurs de séquençage comme étant de véritables variations génétiques de la population virale. Ce seuil de 7% a été établi sur la base d'études comparatives mettant en évidence les capacités de différents « *variant caller* » à identifier correctement des variants génétiques. VVV a été testé sur des données NGS provenant du séquençage d'un isolat de métapneumovirus aviaire (AMPV), ainsi que d'un AMPV recombinant produit par génétique inverse et dont le génome est une copie exacte de l'isolat. À la position 2932 du génome de l'isolat d'AMPV, deux variants génétiques étaient présents à un ratio de près de 50:50. Pour la construction du génome recombinant, l'un des deux variants a délibérément été introduit à cette position. Après le traitement des données de séquençage, VVV a été en mesure de recréer les génomes consensus viraux et de déterminer la composition des populations virales. Un total de 18 et 13 variants génétiques ont été détectés pour l'isolat et le virus recombinant respectivement et comme attendu, un seul nucléotide a été identifié en position 2932 du virus recombinant. Ces données ont été utilisées afin de valider le fonctionnement de VVV.

VVV a également été testé sur des données de séquençage issues d'une souche de coronavirus de la diarrhée épidémique porcine (PEDV), d'une souche d'IBV et d'une souche de TCoV. VVV a produit des séquences consensus génomiques correspondant aux séquences préalablement publiées et un total de 10, 28 et 3 variants génétiques ont pu être détectés respectivement.

Ces différents essais ont mis en évidence l'importance de la quantité de données NGS produites (reflétée par la profondeur de couverture, c'est-à-dire le nombre de lectures pour chaque position de génome, Figure 12) afin de pouvoir distinguer un variant génétique d'une erreur de séquençage. En effet, les erreurs de séquençage surviennent le plus souvent aléatoirement, avec quelques biais de séquence en fonction des technologies utilisées (région homopolymérique pour la technologie IonTorrent, substitutions en troisième position du motif GGCGGG pour la technologie ILLUMINA (Quail et al., 2012)). À l'inverse, un véritable variant génétique présent à 10% dans une population virale sera normalement retrouvé dans 10% des lectures. Par conséquent, obtenir une profondeur de couverture suffisante permet de discriminer correctement un variant génétique d'une erreur de séquençage. Par ailleurs, étant donné que les outils de recherche de variants ne sont pas testés pour des profondeurs de couverture inférieures à 500, il est logique de considérer cette valeur comme étant un seuil minimal afin de procéder à des analyses de populations virales. Cependant, lors du séquençage de génome viraux il est habituel de constater des variations importantes de profondeurs de couverture pour chaque position du génome. C'est pourquoi dans les études présentées ci-après, il a été décidé d'appliquer une médiane de couverture

de 500 lectures NGS avec un minimum de 200 lectures à toutes les positions génomiques. Ce critère étant primordial pour une analyse de variants pertinente, il était important qu'un graphique schématisant la profondeur de couverture pour chaque position fasse partie des résultats de VVV.

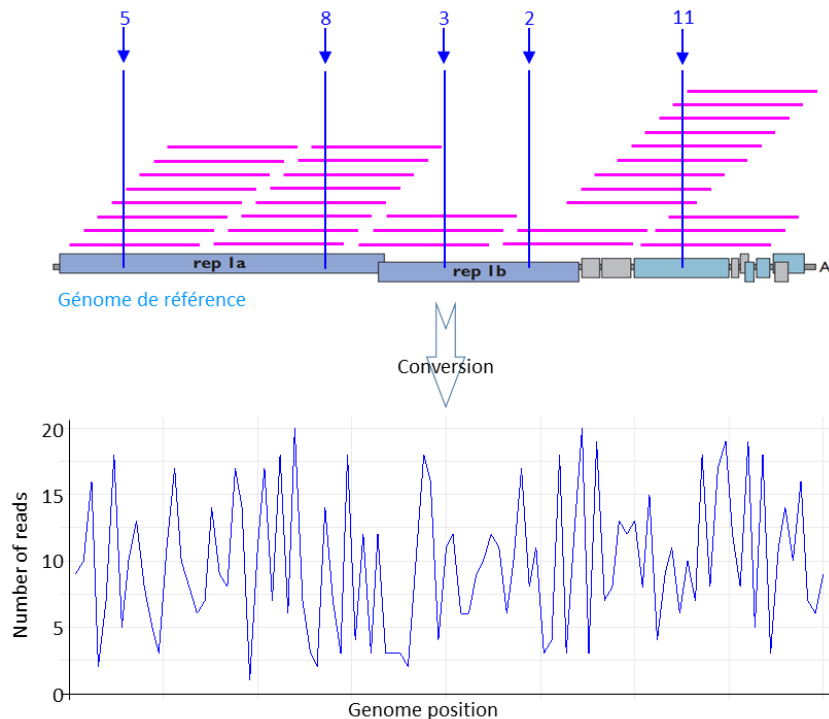


Figure 12: Représentation schématique de la notion de profondeur de couverture.

Le schéma du haut représente un alignement des lectures appelées « reads » (en rose) sur un génome de référence. Certaines positions de génomes sont représentées par des barres verticales bleues. La profondeur de couverture à ces positions est écrite au-dessus des barres bleues et correspond au nombre d'intersection des barres bleues et roses. Le schéma du bas représente le graphique qui est établi après avoir calculé la profondeur de couverture pour toutes les positions du génome.

Ces différents essais ont permis de démontrer les capacités de VVV à gérer l'intégralité du processus de traitement des données, de façon à pouvoir traiter l'ensemble des échantillons de la même façon, et ainsi permettre la détection de variants génétiques composant la population virale. Par ailleurs, la simplicité d'utilisation de VVV pourra aider des biologistes non-bioinformaticiens à approcher les thématiques d'études portant sur les populations virales et l'évolution de celle-ci.

Ce travail a fait l'objet d'une publication dans *Virus Research*.



Viral variant visualizer (VVV): A novel bioinformatic tool for rapid and simple visualization of viral genetic diversity

Alexandre Flageul^a, Pierrick Lucas^b, Edouard Hirchaud^b, Fabrice Touzain^b, Yannick Blanchard^b, Nicolas Etteradossi^a, Paul Brown^a, Béatrice Grasland, Corresponding Author^{a,*}

^a Agence Nationale de Sécurité Sanitaire, de l'environnement et du travail (ANSES) Laboratory of Ploufragan-Plouzané-Niort, Virology, Immunology and Parasitology in Poultry and Rabbit (VIPAC) Unit, Université Bretagne Loire (UBL), France

^b Agence Nationale de Sécurité Sanitaire, de l'environnement et du travail (ANSES), Laboratory of Ploufragan-Plouzané-Niort, Viral Genetic and Biosafety (GVB) Unit, France

ARTICLE INFO

Keywords:

Bioinformatic pipeline
Viral population
NGS

ABSTRACT

Here a bioinformatic pipeline VVV has been developed to analyse viral populations in a given sample from Next Generation Sequencing (NGS) data. To date, handling large amounts of data from NGS requires the expertise of bioinformaticians, both for data processing and result analysis. Consequently, VVV was designed to help non-bioinformaticians to perform these tasks. By providing only the NGS data file, the developed pipeline generated consensus sequences and determined the composition of the viral population for an avian Metapneumovirus (AMPV) and three different animal coronaviruses (Porcine Epidemic Diarrhea Virus (PEDV), Turkey Coronavirus (TCoV) and Infectious Bronchitis Virus (IBV)). In all cases, the pipeline produced viral consensus genomes corresponding to known consensus sequence and made it possible to highlight the presence of viral genetic variants through a single graphic representation. The method was validated by comparing the viral populations of an AMPV field sample, and of a copy of this virus produced from a DNA clone. VVV demonstrated that the cloned virus population was homogeneous (as designed) at position 2932 where the wild-type virus demonstrated two variant populations at a ratio of almost 50:50. A total of 18, 10, 3 and 28, viral genetic variants were detected for AMPV, PEDV, TCoV and IBV respectively. The simplicity of this pipeline makes the study of viral genetic variants more accessible to a wide variety of biologists, which should ultimately increase the rate of understanding of the mechanisms of viral genetic evolution.

1. Introduction

Routine sequencing of nucleic acids molecules started in the late 1970's, using the Sanger's chain-termination method (Sanger et al., 1977) in which different length amplicons are generated using target molecule specific primers and modified nucleotides that randomly terminate elongation. The rise of this technique allowed genome sequencing of multiple living species (Morales and Góes, 2016; Vogel, 2000; Celniker, 2000; Consortium, 1998; Goffeau et al., 1996) but these projects were expensive and time-consuming.

Most recently, in virology and other fields, there has been a substantial increase in the use of a newly developed technique known as high throughput sequencing and globally named as Next-Generation Sequencing (NGS). This has had a positive impact on characterising

genomes of many important viruses and other pathogens. NGS techniques are based on the simultaneous random sequencing of millions of small DNA/RNA fragments and thus much more information about the genomic composition of the sample is obtainable compared to target specific Sanger sequencing where only the target is amplified. In the interest of pathobiology, this ultimately increase the chance of identifying previously unknown organisms, or at least their genetic make-up. What's more, these techniques are rapid and now very cost effective.

One field of study that has increasing interest and takes advantage of NGS is genetic evolution biology. Here, the enormous amount of data that are generated by NGS, allow the genetic evolution biologist to analyse genomic sequences and focus on the number of variants around the majority sequence (consensus) in a given sample. In virology, this is known as analysing the 'viral population' or 'mutant swarm' also known

* Corresponding author.

E-mail address: beatrice.grasland@anses.fr (B. Grasland).

<https://doi.org/10.1016/j.virusres.2020.198201>

Received 15 May 2020; Received in revised form 13 September 2020; Accepted 15 October 2020

Available online 17 October 2020

0168-1702/© 2020 Elsevier B.V. All rights reserved.

as ‘viral quasi-species’ (Domingo et al., 2012; Andino and Domingo, 2015). Observing the totality of the viral genetic population in this way means that it is possible to get a deeper insight about the evolutionary processes that may lead to a change of consensus sequence. Such analysis could be a point of interest for early detection of emerging variant viruses resisting antiviral treatment (Wensing et al., 2019).

NGS data are analysed by different bioinformatics tools that process the enormous amount of data to generate complete genomic sequences. Tools in this process concerning the analysis of viral genetic population are designated as ‘variant callers’. Some popular ‘variant callers’ are VarDict (Lai et al., 2016), Varscan2 (Koboldt et al., 2012) or LoFreq (Wilm et al., 2012). However, to date none of these provides a simple graphic output for the user, which makes data analysis harder as the number of genetic variant increases. For this reason, a bioinformatic tool that is turning NGS raw data to simple graphic outputs (a complete analysis ‘pipeline’) is needed.

This work presents and discusses the development and application of a new bioinformatic pipeline named Viral Variant Visualizer (VVV). VVV produces a simple graphic showing genome locations of genetic variants and their frequencies in the viral genome population, together with a graphical representation of coverage depth of reads aligned on the viral reference; a Krona (Ondov et al., 2011) chart and a detailed table summarizing the precise location of the different variants.

The pipeline was applied to genetic data obtained by NGS from four viruses of veterinary importance: Avian Metapneumovirus (AMPV, *Pneumoviridae*); and three animal coronaviruses: Porcine Epidemic Diarrhea Virus, Turkey Coronavirus and Infectious Bronchitis Virus (PEDV, TCoV and IBV respectively, three members of the *Coronaviridae* family).

2. Methods

Unless specified, in all cases where kits were used, manufacturer’s recommendations were followed.

2.1. Viruses

Two AMPV strains were used to validate the pipeline: a field strain and a copy of this virus produced from a cloned plasmid DNA (Szerman et al., 2019). The field type virus was known to have a mixed population at location 2932 (nt genome position) at a ratio of 50:50, whereas the same position in the cloned virus was engineered in order to be composed of a unique nucleotide which should be evident to a ‘variant calling’ bioinformatics tool. AMPV-C Muscovy duck strain 99178 (wt-AMPV, Genbank accession number [GBAccN]: HG934338.1) had been previously isolated in Vero cells (Toquin et al., 1999) and the cloned copy of this virus (rc-AMPV) had been previously produced from cloned plasmid DNA (Szerman et al., 2019) (Table 1).

Three coronaviruses were chosen due to their notorious capacity for viral diversity. PEDV/USA/2014/Iowa (GBAccN: MF373643.1) was a field virus isolated from jejunum of PEDV-affected piglets in USA

(Gallien et al., 2018). IBV-Mass41 and European TCoV 080385d (unpublished and GBAccN: KR822424.1 respectively) strains were harvested from allantoic fluid and intestines of 23 day old embryos respectively according to previously described methods (Coronaviruses: Methods and Protocols, 2015; Guionie et al., 2013).

2.2. RNA extraction

RNA was extracted from IBV-Mass41 using RNeasy® kit (Qiagen). RNAs from TCoV-080385d, wt-AMPV and rc-AMPV were extracted using TRIzol™ LS reagent (Invitrogen), following manufacturer’s instructions. RNA was extracted from PEDV with a modified TRIzol™ (Invitrogen) protocol as follows: 250 µL of sample was mixed with 1 mL of TRIzol™ and 200 µL of chloroform/isoamyl-alcohol (24:1). The aqueous phase containing RNA was harvested then mixed with 300 µL of TRIzol™ and 60 µL of chloroform/isoamyl alcohol (24:1). The new aqueous phase was harvested and RNA was precipitated with 1 µL of glycogen and 150 µL of cold isopropanol. After 15 min incubation at –20 °C, a centrifugation step was performed at 12,000 g for 10 min. The resulting pellet was washed twice with 75 % ethanol and finally resuspended in 50 µL of RNase-free water. This constituted PEDV extracted RNA.

2.3. Library preparation

Two different NGS technologies were used Ion Torrent and Illumina in this project and thus two different RNA library preparations were carried out.

i) Ion Torrent sequencing: a maximum concentration of RNA was used respecting an upper ceiling of 500 ng in a maximum volume of 20 µL. This was treated with 4 units of DNase (Turbo DNA-free™ kit from Life Technologies). Following this treatment, ribosomal RNAs were removed (ribodepletion) using Low Input RiboMinus™ Eukaryote System v2 (Life Technologies). RNA libraries were then prepared with Ion Total RNA-seq kit 2 (Life Technologies). Briefly, this process involved enzymatic fragmentation of RNA and ligation of adaptors to the fragmented RNA to generate cDNA from RNA molecules by reverse transcription. Resulting cDNAs were then amplified and barcoded for identification. Finally, library quality was checked on a Bioanalyzer from Agilent (High Sensitivity DNA Kit and RNA 6000 Pico Kit, Agilent).

ii) Illumina sequencing: RNA library was prepared as described previously (Brown et al., 2016). Briefly, after extraction, RNA was treated with 4 units of DNase (Turbo DNA-free™ kit from Life Technologies) and purified (RNeasy® MinElute® Cleanup kit; Qiagen). Size distribution of RNA fragments was checked on Agilent 21000 Bioanalyses (LabChip RNA 6000 Nano kit). cDNAs were generated and amplified using Ovation® RNA-Seq System V2 (NuGEN). The library was constructed with Encore® Rapid DR Multiplex System1-8 (NuGEN), using 1 µg of cDNA to obtain fragments of 150 to 500bp.

Table 1
Viral sample features.

Virus name	Strain	Genbank accession number	Origin	Method of Extraction	NGS technology
wt-AMPV	AMPV-C - Muscovy_duck/1999/99178/ France	HG934338.1	Viral stock prepared on Vero cells after 1 passage	TRIZol™ (Invitrogen)	Ion Torrent
rc-AMPV	Engineered viral clone of AMPV-C - Muscovy_duck/1999/99178/ France	unpublished	Viral stock prepared on Vero cells	TRIZol™ (Invitrogen)	Ion Torrent
PEDV	PEDV/USA/2014/Iowa	MF373643.1	Intestinal tissue homogenate from PED affected piglets	TRIZol™ (Invitrogen)	Ion Torrent
TCoV	European turkey coronavirus 080385d	KR822424.1	Viral Stock prepared on eggs after 6 passages	TRIZol™ LS (Invitrogen)	Illumina MiSeq
IBV	Mass41	unpublished	Viral Stock prepared on eggs after 7 passages	RNeasy™ kit (Qiagen)	Ion Torrent

2.4. VVV development

The VVV pipeline was implemented on an Ubuntu server 16.04 × 86.64 with 28 cores Intel(R) Xeon(R) 2,60 GHz, 58 threads and 378Go of memory and can be used freely via the publically available Galaxy Instance of ANSES (<https://viralvariant.anses.fr>). Also, source code of the VVV can be found here: https://github.com/ALFLAG/Viral_Variant_Visualiser.

The complete VVV was developed using Snakemake (Köster and Rahmann, 2012) (v5.2.2), which is a management system that helps the development of workflows for bioinformatics data analysis. In this pipeline, many different bioinformatic tools (described below) were used as well as some in-house Python3, R and Perl scripts. All tools were installed using the package and environment management system Conda (Computeur Software, 2018) (v4.6.8). Unless specified, all tools were used with the default parameters.

Briefly, raw data are trimmed with Trimmomatic (Bolger et al., 2014) (v0.38) to remove low quality reads and if paired-end data are provided, they are merged as single-end data using Pear (Zhang et al.,

2014) (v0.9.11). The metagenomic composition of the sample is determined using Megablast (Morgulis et al., 2008) (v2.7.1), Krona and in-house script on NCBI's nucleotide database. If a reference sequence is provided with the NGS data, VVV automatically aligns trimmed reads on the provided sequence using BWA (Li, 2013) (v0.7.17-r1188) before performing the variant calling with the Java version of VarDict (Lai et al., 2016) (v1.5.8). If no reference sequence is provided, VVV generates a consensus viral sequence. This process starts by selecting reads that align to a custom database made with all viral sequences from GenBank, using Megablast. The most represented complete viral genome is used to align selected reads using BWA and the median coverage depth is calculated. Then, adequate subsampling is performed with SeqTK (<https://github.com/lh3/seqtk>) (v1.3) to obtain sufficient amount of reads that generate a median coverage depth of 60, required for *de novo* assembling using SPAdes (Bankevich et al., 2012) (v3.12) and Mira (Chevreux et al., 1999) (v4.9.6). Generated contigs are aligned with MegaBlast on NCBI's nucleotide database to identify the best viral match that is used to determine the consensus genome of the viral sample (using Samtools (Li et al., 2009) mpileup command [v1.9.9] with quality

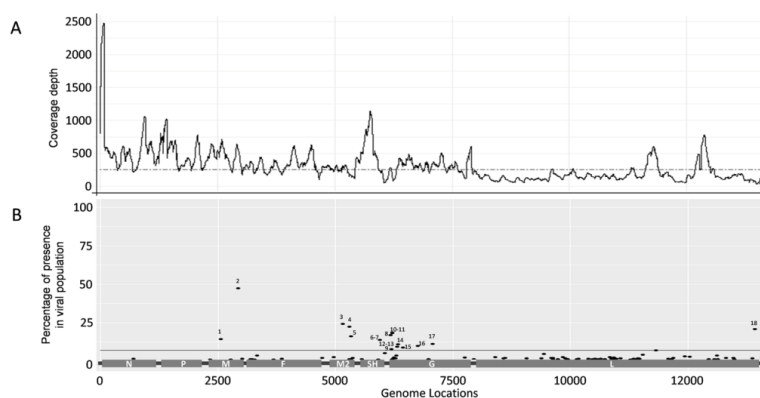
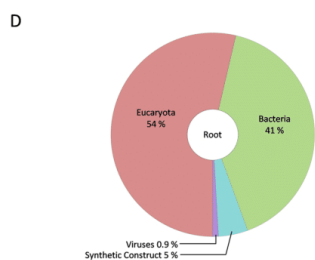


Fig. 1. Output results of VVV pipeline obtained for wt-AMPV. A) graphical representation of the number of reads per genome location, dotted line represents the median coverage depth ; B) viral genetic variants map displaying the detected genetic variants (black dots) present in the viral population above the user defined threshold cut-off value (solid horizontal line 7%). Numbers above the black dots are indices that identify the variant in the graphic and must be reported in C) the detailed table which informs on the position, type of variation, sequences present in the 5' and 3' side of the reference nucleotides, and a warning flag that is turned to 'yes' to indicate if the variant is detected within a homopolymer region (repetition of at least three identical nucleotides) ; D) a krona chart displaying the metagenomic composition of the sample.

Index	Location	Ref	Alt	Frequency	Gene	5' seq	3' seq	Homopolymer*
1	2565	T	C	0.1426	M	TGCTCTAAGTTTGTCAAC	CAGCCAAGGCAGTCGGAAAG	no
2	2932	G	T	0.4729	M	GAGATATGCTCGAAGTCGA	ATAAACCCAAAGGCCACACC	no
3	5152	TA	T	0.2413	M2	GAAGCTCCGAAAGGAAAAC	AAAAAATGGCAAGCTGAT	yes
4	5301	TA	T	0.2228	M2	TTAGATGATATGGTATGGAC	AAAAATGATGGTAGACATA	yes
5	5321	T	C	0.1609	M2	TAAAATGAAATGGTAGACA	AAATTCACGAAATAGTATA	no
6	5939	T	C	0.1377	SH	GTGCATGCATATATGCGAG	AAATGGGATTTCTGGTARG	no
7	5949	T	C	0.1387	SH	ATATATGCAGTAAATTGGGA	TGTGTTATGCTTGAGAGAG	no
8	6165	T	C	0.1681	G	AGTCAACATGAGAGTCAAGA	AGAGAAATGTGGCAAGTCAC	no
9	6188	G	T	0.0784	G	AGAATGTTGGCAAGTCACAA	AGCTTAGATGTTAAAGTCAAG	no
10	6199	T	C	0.1818	G	AAGTCAACAAGCTTAGAGT	AAAGTCAAGAATTTTATAAA	no
11	6211	TTTTAT	CCCTAC	0.1835	G	CTTAGAGTAAAGTCAAGAA	AAAAAGCTGATTTGCAAGA	no
12	6308	T	C	0.095	G	ATATAATGTTGCTGTCATG	ATGTGGAGTCAAATAAGGCC	no
13	6317	T	C	0.0951	G	TGTCGTGTCATGATGTCGAG	CAATAAGGCCCTGAGCTTA	no
14	6322	T	C	0.1094	G	GTCATGATGTCGAGTCAAA	AAGCCCTGAGCTTATGCGAG	no
15	6438	T	C	0.0872	G	AACGACCAACAACAGGAGCA	AGAAGCAAGGGGGACAAAA	no
16	6745	T	C	0.0962	G	ACCAAGAAAGCCGACAGCA	ACCAAGCAACAGGCAAGC	no
17	7071	T	C	0.1101	G	AGAAACCAACAAGAGGACCA	AACAACAAGGGCCCAAAA	no
18	13923	TA	T	0.2105	L	AATCTGGCAGTCTGCAACT	AAAAGCTAATTAAGTGACA	yes

*NB: an homopolymer region is set to 'yes' if there is a succession of at least 3 identical nucleotides.



values set to 0, and in-house python script). Finally, all trimmed reads are aligned against the consensus sequence with BWA. The aligned reads are then compared to the NCBI nt database to look for non-viral match, as they might be the results of chimeric reads, i.e. reads with both viral and non-viral sequences that can be generated during NGS library preparation due to the chemistry of the ligation step. These chimeric reads are removed during a second alignment against the matched non-viral sequences. Remaining reads are then aligned on the consensus genome and a variant calling step is performed with the Java version of VarDict (Lai et al., 2016). Viral genetic variants are visualised on a graphic showing the genome organisation on the X-axis and the percentage of genetic variants for each position on the Y-axis. Genetic variants are considered as 'true' when present above a user defined threshold and set in the present study as 7%.

3. Results

3.1. VVV output graphics

Fig. 1 displays results from VVV output after processing wt-AMPV.

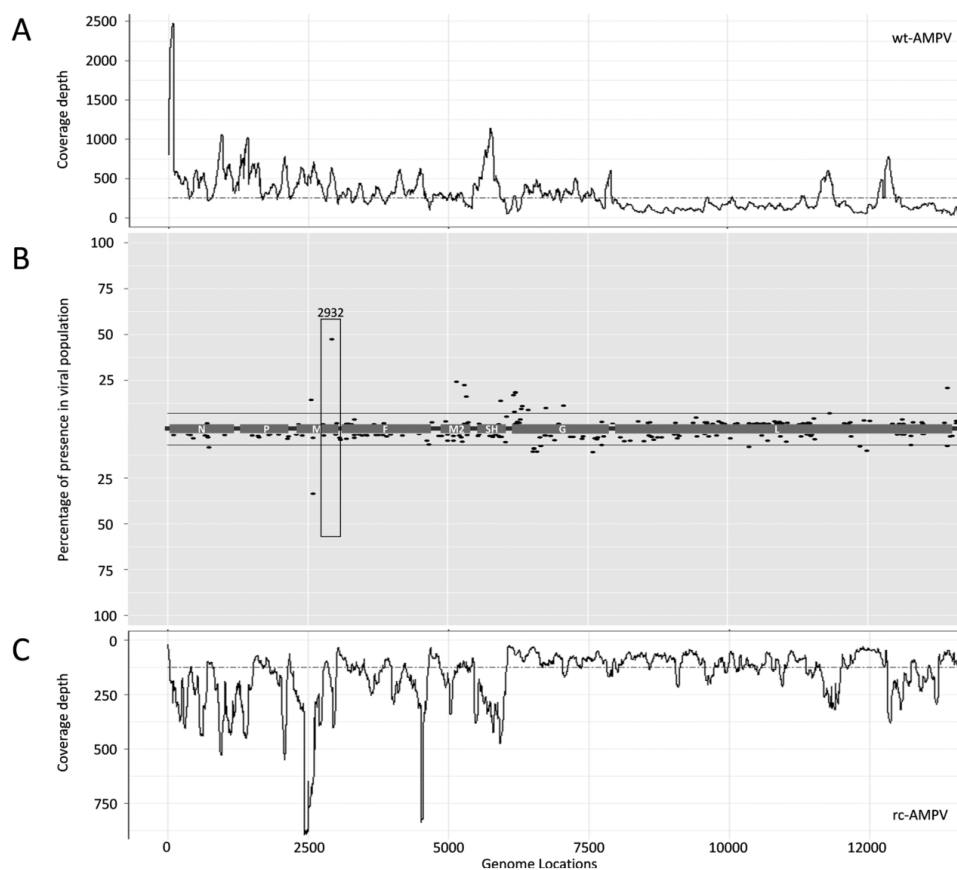


Fig. 2. Viral genetic variant map and genome coverage for wild-type (wt) and recombinant (rc) AMPV. A and C) White graphics wt-AMPV and rc-AMPV show the coverage depth along the two genomes. Dashed lines show median coverage depth (257 for wt-AMPV and 125 for rc-AMPV). B) Grey graphics show viral genetic variants (black dots) and the gene locations (N, P, M, F, M2, SH, G, L) for wt-AMPV and rc-AMPV. Legend on the y-axis defines the percentage at which each variants was present. Solid horizontal lines delimitate the variant cut-off, here set to 7%. Vertical box highlights location 2932 for both viruses. In graphic wt-AMPV a variant at 47% can be seen which is absent in graphic rc-AMPV.

First, a representation of the number of reads per genome location with the median coverage depth are showed (Fig. 1A) followed by a viral genetic variants map (Fig. 1B). A table gives more detail about the variants with a frequency above 7%, the cut-off threshold value used in this study (Fig. 1C), and finally a krona chart displays the metagenomic composition of the sample (Fig. 1D).

3.2. Validation of VVV with wild-type and recombinant AMPV

VVV was first tested for its validity on a sequence data set obtained from RNA extracted from wt-AMPV and rc-AMPV for the reasons described in section "Viruses" above. Briefly, wt-AMPV was known to have a mixed population of sequences displaying either a G or a T base at genome position 2932 at ratio of 50:50, whereas rc-AMPV was known to display a G base, in 100% of its genome population. After processing wt-AMPV Ion Torrent raw data with VVV, a consensus sequence identical to the previously published sequence, was determined using 50927 AMPV reads (from a total of 8024707 single-end reads with a mean size of 87 nucleotides). From these data, the median number of reads per genome position was 257 (Fig. 2A). Also, VVV identified 18 genome positions

[2565, 2932, 5152, 5301, 5321, 5939, 5949, 6165, 6188, 6199, 6211, 6308, 6317, 6322, 6438, 6745, 7071, 13923] where variants existed above the chosen threshold and these were mostly located in the region coding for the hypervariable attachment glycoprotein (Fig. 2B, G gene). At position 2932, where a ratio of 50:50 G:T was expected, VVV demonstrated 53 % G and 47 % T base. The variant present at 47 % can be seen for wt-AMPV in Fig. 2B in the black box. After processing rc-AMPV Ion Torrent raw data with VVV, an exact copy of the wt-AMPV genome consensus was found, determined using 30873 AMPV reads (from a total of 726377 single-end reads with a mean size of 83 nucleotides) and a median number of reads per genome position of 125 (Fig. 2C). Thirteen viral genetic variants were found dispersed along the genome [positions 727, 2596, 6525, 6531, 6538, 6591, 6660, 7585, 10381, 12364, 12480, 13282, 13923] and not concentrated around the hypervariable attachment glycoprotein coding sequence (Fig. 2B). As expected no viral genetic variant was detected at position 2932 (Fig. 2B, black box rc-AMPV).

3.3. VVV assessment of PEDV, TCoV and IBV

After processing PEDV/USA/2014/Iowa NGS raw data with VVV, a consensus sequence that matched the previously published sequence (Gallien et al., 2018) was determined using 87744 PEDV Ion Torrent reads (from a total of 12267490 single ends reads with a mean size of 100 nucleotides) and a median number of 209 reads per genome position. Ten viral genetic variants were found along the genome, seven

were located in the orf1ab gene [positions 2383, 7845, 7888, 8708, 9337, 10532, 17077], one in the S gene [position 22696] and two in ORF3 gene [positions 24964, 25102] (Fig. 3).

After processing the European TCoV-080385d consensus genome and Illumina raw data with VVV, it was observed that 31272 TCoV reads (from a total of 2253856 paired-end reads with a mean size of 251 nucleotides) were aligned on the reference genome, given a median number of 64 reads per genome position. Three viral genetic variants were found, two of them were located in the orf1ab gene [positions 13495, 14546] and the third was located in the 5b gene [position 25900] (Fig. 4).

After processing IBV-Mass41 Ion Torrent raw data with VVV, a consensus sequence was determined using 229752 IBV reads (from a total of 436755 single-end reads with a mean size of 75 nucleotides) and a median number of 542 reads per genome position (Fig. 5). Twenty-eight viral genetic variants were found dispersed along the genome [positions 73, 2439, 4304, 5261, 6347, 6722, 6869*, 7082*, 7977, 8832, 9079*, 11088, 11568*, 13771*, 15082, 15084, 15157*, 16049, 17512, 20029, 20440, 21541, 23838, 25245, 25955*, 26848*, 26916*, 26955*]. Variants marked with an asterisk were found at more than 40 % of presence in the viral population.

4. Discussion

In virology, genetic variants are those that contain nucleotides that differ from the consensus sequence of the population at any given

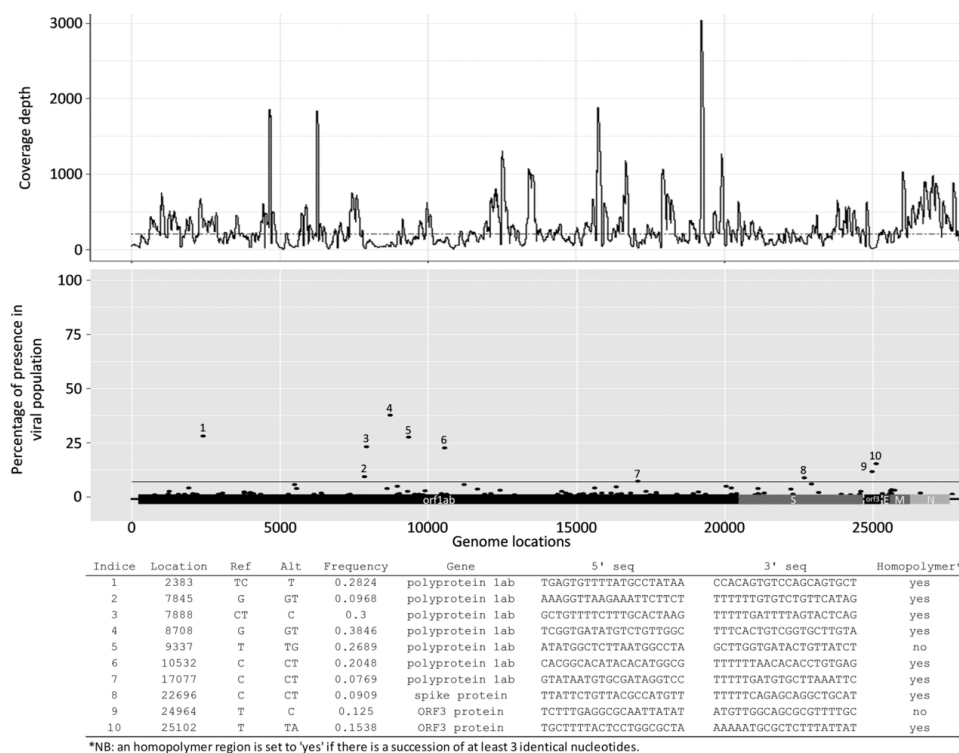


Fig. 3. Viral genetic variant map and genome coverage for Porcine Epidemic Diarrhea Virus (PEDV). The grey graphic shows viral genetic variant map of PEDV and the gene locations (orf1ab, S, orf3, E, M, N). Legend on the y-axis define the percentage at which each variants was present. Solid horizontal line delimitate the variant cut-off. The white graphic shows the coverage depth of the genome and the broken line shows median coverage of 209 for PEDV. Table below gives details about the detected genetic variants.

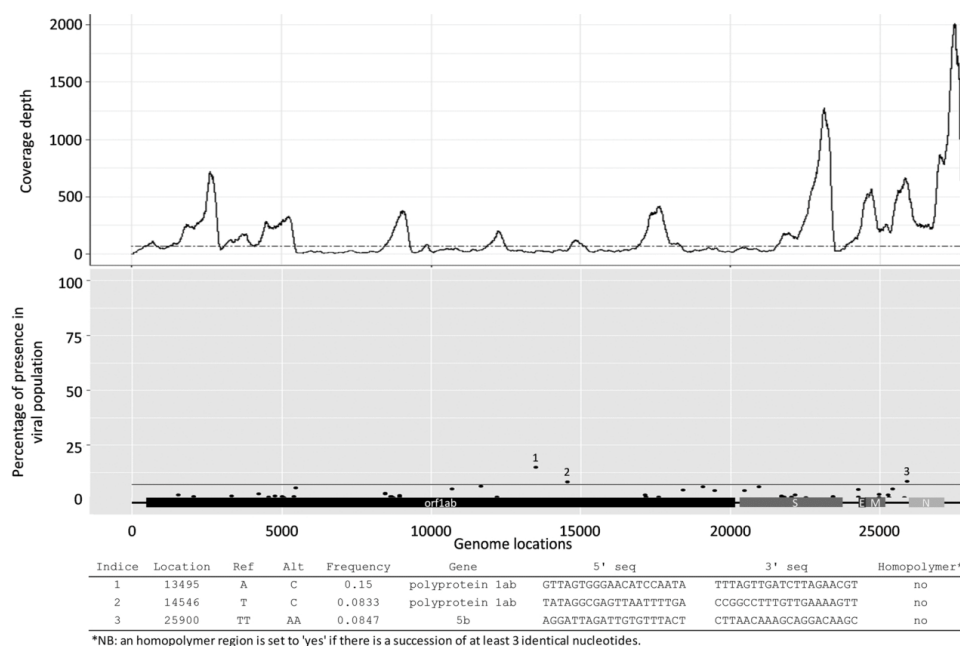


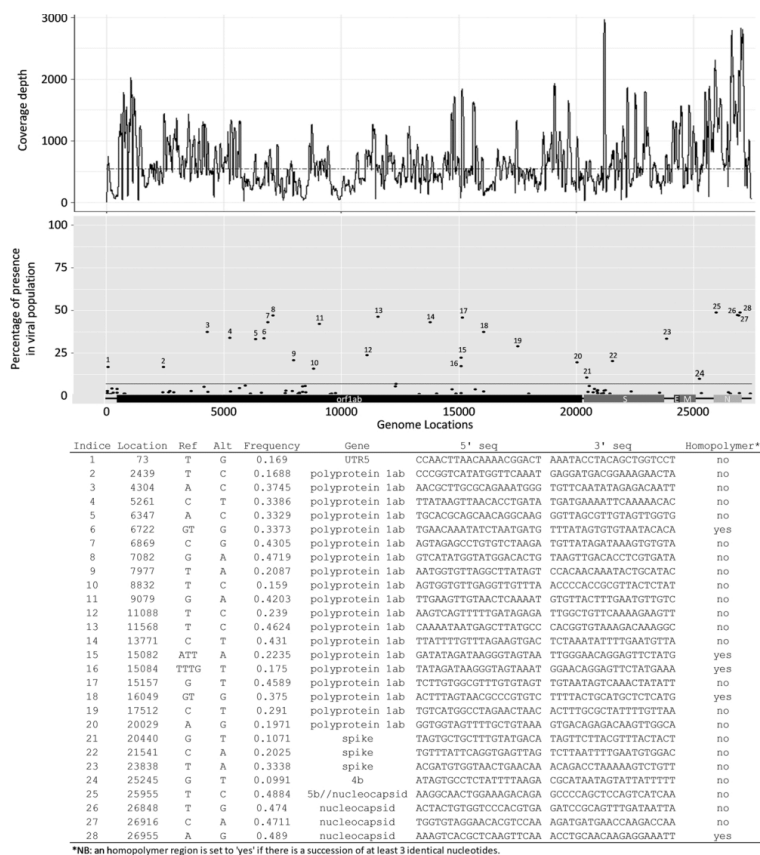
Fig. 4. Viral genetic variant map and genome coverage for Turkey Coronavirus (TCoV). The grey graphic shows viral genetic variant map of TCoV and the gene locations (orf1ab, S, E, M, N). Legend on the y-axis define the percentage at which each variants was present. Solid horizontal line delimitates the variant cut-off. The white graphics show the coverage depth of the genome and the broken line shows median coverage of 66 for TCoV. Table below gives details about the detected genetic variants.

location. Studying genetic variants provides information on all genomic molecules present in a viral population and is important for understanding the interplay between the different genomes of the population during any evolutionary process. Studying only variations at the consensus level does not provide information on how the population arrived at a particular sequence majority. Furthermore, consensus sequences show only the most probable nucleotide at any given position and may even not exist as a single molecule in the actual viral population, as they are generated from fragments of viral RNA and not from full-length genomes. Here, the VVV bioinformatic pipeline was specifically designed to detect viral genetic variants in a viral genetic population using a recent variant calling program, VarDict (Lai et al., 2016). This program was chosen based on independent study that compared the efficacy of different variant calling tools, and in which VarDict showed the best performance in term of sensitivity (Sandmann et al., 2017). However, it must be noted that when multiple minority variants are located at the same position, VarDict only indicates the one with the highest frequency.

Although NGS technologies are prone to errors (mostly insertion/deletion errors for IonTorrent, with an important bias in homopolymer tracts; mostly transition error [C > T] in Illumina) (Quail et al., 2012), read error correction programs were not incorporated in VVV as they tend to categorise real genetic variant as errors (Song et al., 2017). To avoid categorizing sequencing errors as genetic variants, and because variant callers have reduced capacity to detect low frequency genetic variant (below 5% of presence in the viral population) (Sandmann et al., 2017), a cut-off variant threshold of 7% was used instead in the analysis, which is higher than in other studies (Dinis et al., 2016; Arzt et al., 2019). Future improvements of this tool could be the addition of a second variant calling tool as has been suggested before (Said Mohamed et al., 2018). However, doing so, more computational time and

resources would be required, and also, variations in cut-off value thresholds, depending on the variant caller used (Sandmann et al., 2017).

Consensus sequences and viral population compositions were obtained from raw NGS data of AMPV, PEDV, TCoV, IBV and were represented in a simple graphic when processed by VVV. For rc-AMPV, a low number of genetic variants were observed, which was expected due to the clonal origin of this virus however, the variants that were present highlight the capacity of viral RNA polymerase to incorporate erroneous nucleotides. Importantly, for the validation of VVV, the rc-AMPV population was homogenous at position 2932, where the parental wt-AMPV was heterogeneous at the same location. Different percentages of variants were observed above the chosen threshold for the different viruses, with low number of genetic variants detected for PEDV and TCoV. In contrast, IBV demonstrated the most important number of genetic variants and may reflect why there are so many different "genotypes" or strains reported for this virus (Valastro et al., 2016). The IBV sample used in the current study had been prepared by multiple passages on embryonated chicken eggs and thus the high number of genetic variants could have been generated during this process. Contrary to this hypothesis is the fact that the TCoV sample, also a viral stock produced by multiple passage in embryonated eggs, showed a limited number of variants. One possible explanation for this could be related to the amount of NGS data obtained. For TCoV the number of reads was much lower than that of IBV as reflected by the coverage depth for each sample (IBV [min = 0, max = 2968, median 542], TCoV [min = 0, max = 1944, median = 64]), and lower coverage depth reduces the capabilities of detecting true variants. Such results, highlight the importance of having sufficient coverage depth when performing variant calling analysis: a median coverage of 500, that is set for evaluating variant calling tools (Sandmann et al., 2017) and at least 200 at all positions, because of the



*NB: an homopolymer region is set to 'yes' if there is a succession of at least 3 identical nucleotides.

Fig. 5. Viral genetic variant map and genome coverage for Mass41 Infectious Bronchitis Virus (IBV) strain. Grey graphic shows viral genetic variants (black dots) and the gene locations (orf1ab, S, E, M, N) without accessory genes locations. Legend on the y-axis define the percentage at which each variants was present. Solid horizontal lines delimitate the variant cut-off. White graphic shows the coverage depth along IBV genome. Broken line shows median coverage depth of 542.

heterogeneous coverage depth of viral genomes, should be the minimum.

In addition to analysing the variants within a particular sample, the variant map was used to check the consensus sequence, by identifying consensus errors displayed as variants above 50 %. Finally, the fact that VVV can accept a reference sequence as input together with the raw NGS data, as it was done for TCoV, provides a means to compare any sequenced viral sample to any related reference genome. Consensus modifications will appear above the 50 % of presence showing potential selection of variants.

To conclude, the simplicity of this pipeline makes studying viral genetic populations and the evolution of these, more accessible to scientists who do not have a bioinformatic background, yet have expertise in their biological field of study. Furthermore, this pipeline is currently being used to address the evolution of Avian Coronavirus after successive passages in vaccinated or unvaccinated animals. Future improvement of this tool would be the capacity to perform the same analysis with the genome of segmented viruses.

Funding

This work was funded by Agence Nationale de Sécurité Sanitaire, de

l'Environnement et du travail (ANSES), the "Région Bretagne", and the "Université Bretagne Loire" (UBL) and "Université de Rennes" (UNIR) (code of convention : 180019).

CRedit authorship contribution statement

Alexandre Flageul: Formal analysis, Software, Methodology, Writing - original draft. **Pierrick Lucas:** Methodology. **Edouard Hirsch:** Software. **Fabrice Touzain:** Software. **Yannick Blanchard:** Resources. **Nicolas Etteradossi:** Project administration. **Paul Brown:** Conceptualization, Funding acquisition, Validation, Writing - review & editing. **Béatrice Grasland:** Conceptualization, Funding acquisition, Validation, Writing - review & editing.

Declaration of Competing Interest

Authors declare no conflict of interest.

Acknowledgment

The authors thank Chantal Allée, Céline Courtillon, Evelyne Lemaitre, Cecilia Bernard, Lionel Bigault, and Nathan Szman for their

assistance in providing the viral samples.

References

- Andino, R., Domingo, E., 2015. Viral quasispecies. *Virology* 479–480, 46–51.
- Arzt, J., et al., 2019. The evolution of a super-swarm of foot-and-mouth disease virus in cattle. *PLoS One* 14.
- Bankevich, A., et al., 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477.
- Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120.
- Brown, P.A., et al., 2016. First complete genome sequence of European turkey coronavirus suggests complex recombination history related with US turkey and guinea fowl coronaviruses. *J. Gen. Virol.* 97, 110–120.
- Celniker, S.E., 2000. The *Drosophila* genome. *Curr. Opin. Genet. Dev.* 10, 612–616.
- Chevreaux, B., Wetter, T., Suhai, S., 1999. Proceedings of the German Conference on Bioinformatics, GCB'99 Genome sequence assembly using trace signals and additional sequence information. *Comput. Sci. Biol.* 45–56.
- Computer Software, 2018. *Anaconda Software Distribution*.
- Consortium, 1998. *T. C. elegans S. Genome Sequence of the Nematode C. elegans: A Platform for Investigating Biology*. *Science* 282, 2012–2018.
- Coronaviruses: Methods and Protocols.** Humana Press, 2015.
- Dinis, J.M., et al., 2016. Deep sequencing reveals potential antigenic variants at low frequencies in influenza A virus-infected humans. *J. Virol.* 90, 3355–3365.
- Domingo, E., Sheldon, J., Perales, C., 2012. Viral quasispecies evolution. *Microbiol. Mol. Biol. Rev.* 76, 159–216.
- Gallien, S., et al., 2018. Better horizontal transmission of a US non-InDel strain compared with a French InDel strain of porcine epidemic diarrhoea virus. *Transbound. Emerg. Dis.* 65, 1720–1732.
- Goffeau, A., et al., 1996. Life with 6000 genes. *Science* 274, 546–567.
- Guionie, O., et al., 2013. An experimental study of the survival of turkey coronavirus at room temperature and +4°C. *Avian Pathol.* 42, 248–252.
- Koboldt, D.C., et al., 2012. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* 22, 568–576.
- Köster, J., Rahmann, S., 2012. Snakemake—a scalable bioinformatics workflow engine. *Bioinformatics* 28, 2520–2522.
- Lai, Z., et al., 2016. VarDict: a novel and versatile variant caller for next-generation sequencing in cancer research. *Nucleic Acids Res.* 44 e108–e108.
- Li, H., 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *ArXiv13033997 Q-Bio*.
- Li, H., et al., 2009. The sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079.
- Moraes, F., Góes, A., 2016. A decade of human genome project conclusion: scientific diffusion about our genome knowledge. *Biochem. Mol. Biol. Educ.* 44, 215–223.
- Morgulis, A., et al., 2008. Database indexing for production MegaBLAST searches. *Bioinformatics* 24, 1757–1764.
- Ondov, B.D., Bergman, N.H., Phillippy, A.M., 2011. Interactive metagenomic visualization in a Web browser. *BMC Bioinformatics* 12, 385.
- Quail, M.A., et al., 2012. A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics* 13, 341.
- Said Mohammed, K., et al., 2018. Evaluating the performance of tools used to call minority variants from whole genome short-read data. *Wellcome Open Res.* 3.
- Sandmann, S., et al., 2017. Evaluating variant calling tools for non-matched next-generation sequencing data. *Sci. Rep.* 7, 43169.
- Sanger, F., Nicklen, S., Coulson, A.R., 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* 74, 5463–5467.
- Song, L., et al., 2017. Comparison of error correction algorithms for Ion Torrent PGM data: application to hepatitis B virus. *Sci. Rep.* 7, 8106.
- Szerman, N., et al., 2019. The small hydrophobic (SH) gene of North American turkey AMPV-C does not attenuate nor modify host tropism in recombinant European duck AMPV-C. *Virology* 526, 138–145.
- Toquin, D., Bayon-Auboyer, M.H., Eterradosi, N., Jestin, V., Morin, H., 1999. Isolation of a pneumovirus from a Muscovy duck. *Vet. Rec.* 145, 680.
- Valastro, V., et al., 2016. S1 gene-based phylogeny of infectious bronchitis virus: an attempt to harmonize virus classification. *Infect. Genet. Evol. J. Mol. Epidemiol. Evol. Genet. Infect. Dis.* 39, 349–364.
- Vogel, G., Sanger, 2000. Will sequence zebrafish genome. *Science* 290, 1671–1671.
- Wensing, A.M., et al., 2019. Update of the drug resistance mutations in HIV-1. *Top. Antivir. Med.* 27, 111–121.
- Wilm, A., et al., 2012. LoFreq: a sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. *Nucleic Acids Res.* 40, 11189–11201.
- Zhang, J., Kobert, K., Flouri, T., Stamatakis, A., 2014. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* 30, 614–620.

7 Enrichissement des échantillons en ARN viral

Grâce au développement de VVV, les variants génétiques viraux pouvaient être identifiables facilement à partir de données NGS. Malheureusement, les premiers séquençages réalisés à partir d'échantillons de trachées issues de poulets infectés par IBV et à partir de contenus intestinaux de dindes infectés par le TCoV ont révélé un problème important : la capacité à obtenir des quantités suffisantes de lectures issues de génomes viraux pour atteindre une profondeur de couverture suffisante.

Dans un échantillon de tissu, les acides nucléiques du virus à séquencer sont mélangés aux acides nucléiques issus des cellules qui composent le tissu. Grâce au NGS, tous ces acides nucléiques pourront être séquencés. Mais lorsque les acides nucléiques du virus sont présents en quantité minoritaire par rapport aux acides nucléiques eucaryotes, le grand avantage du NGS devient problématique, car il est alors difficile d'obtenir de grandes quantités de données d'origine virale, sans obtenir encore plus de données d'origine eucaryote. Et si l'augmentation de la quantité totale de données produites lors du séquençage pouvait permettre d'augmenter la quantité de données d'origine virale, cela augmenterait également la quantité de données d'origine eucaryote, conduisant à une augmentation du coût du séquençage. Le prix du séquençage devient alors un facteur limitant notamment à cause du grand nombre d'échantillons à séquencer. Un problème d'optimisation se posait donc : comment obtenir le maximum de données provenant des génomes viraux tout en limitant la quantité totale de données produites ? Différentes méthodes ont été envisagées afin de résoudre ce problème.

Une amplification des ARN viraux a d'abord été envisagée via l'utilisation de la RT-PCR. Cette approche présentait l'avantage de permettre une amplification spécifique des ARN d'intérêt afin d'augmenter leurs proportions relatives dans l'échantillon. Malheureusement, bien que souvent utilisé ce type d'approche nécessite l'utilisation d'amorces produites à partir de séquences issues des bases de données. Leur utilisation peut donc entraîner l'apparition d'un biais d'amplification avec une hybridation préférentielle des amorces sur des séquences connues. Par ailleurs, l'utilisation de la PCR présentait également le risque d'incorporation de nucléotides erronés lors des phases d'élongations qui fausseraient les analyses de variants génétiques minoritaires. L'amplification PCR n'a donc pas été retenue.

Une autre approche consistait à purifier les particules virales par l'intermédiaire de l'ultracentrifugation, ou de traiter les tissus biologiques broyés avec des exonucléases spécifiques de l'ARN (RNAses) de sorte à dégrader les ARN non protégés par des particules virales. Ces deux types d'approches permettent de retirer les ARN libérés par les cellules lors du broyage, mais nécessitent l'utilisation d'importante quantité de matériel de départ afin de pallier aux pertes

importantes subies lors du traitement de l'échantillon, ce qui fait que ce type d'approche a également été mis de côté.

En virologie aviaire, une autre approche peut être envisagée : l'inoculation sur œuf. Cette méthode est largement utilisée pour amplifier les stocks viraux et selon la voie d'inoculation elle permet de combiner à la fois l'amplification et la purification partielle des particules virales.

Dans le cas de l'IBV, un protocole de séquençage NGS préconisait d'effectuer une inoculation du virus dans la cavité allantoïdienne d'un œuf de poule embryonné de 9 jours d'âge et de récolter le liquide allantoïdien à partir duquel les ARN viraux étaient extraits (Keep et al., 2015). Deux avantages étaient observés avec cette méthode : i) l'inoculation sur œuf permettait une amplification de tous les génomes viraux, indépendamment de la présence de variants génétiques, grâce à la réplication des virus dans la cavité allantoïdienne ; ii) le liquide allantoïdien était très pauvre en matériel génétique issu de l'hôte, permettant d'avoir très peu de données NGS d'origine eucaryote. Après avoir réalisé un essai en suivant le protocole d'inoculation et d'incubation des œufs mis au point au laboratoire ANSES de Ploufragan (incubation de six jours au lieu d'une journée comme indiqué par Keep et al.), une excellente amplification du virus a été obtenue et pour des quantités similaires de données NGS produites, une meilleure profondeur de couverture du génome viral a été obtenue après un passage sur œuf de poule embryonné (Figure 13).

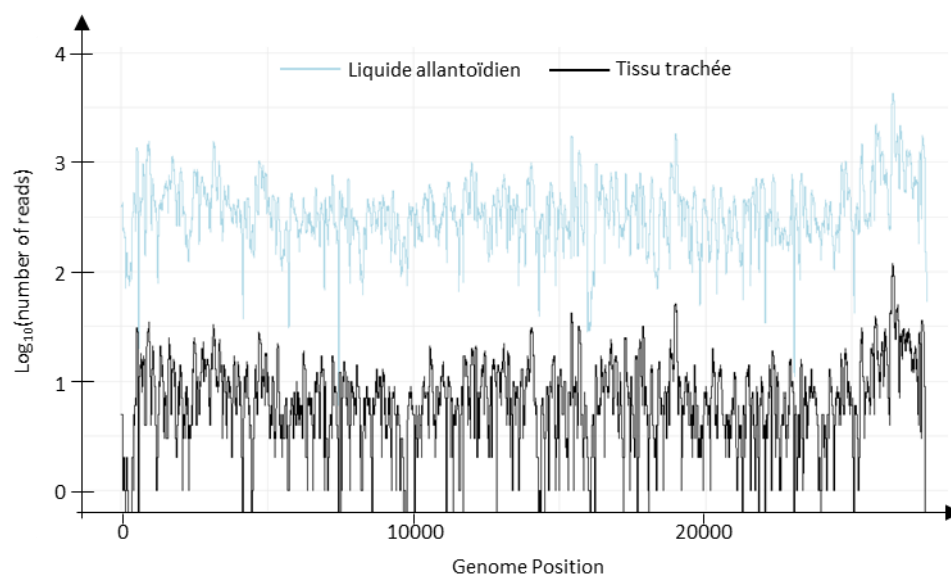


Figure 13: Comparaison des profondeurs de couvertures issues du séquençage IBV, avant et après inoculation sur œuf de poule embryonné. Un total de 247896 lectures a été produit en NGS à partir d'ARN extrait du broyat de trachée infectée par IBV (profondeur de couverture obtenue en noir), et un total de 261549 lectures à partir d'ARN extraits du liquide allantoïdien (profondeur de couverture obtenue en bleu) issu d'un œuf de poulet embryonné inoculé avec le broyat de trachée séquençé auparavant.

Par ailleurs, en plus d'un enrichissement en ARN viral (sans augmenter la quantité d'ARN d'origine eucaryote, qui ne peut pas se répliquer dans l'œuf embryonné), le choix de cette méthode apportait également la preuve que les virus présents dans les organes des poulets infectés étaient réellement infectieux.

Afin de pouvoir réaliser l'analyse des variants génétiques d'IBV après plusieurs passages, il fallait tout d'abord vérifier qu'un passage sur œuf permettait de conserver les populations virales présentes avant le passage sur œuf. En utilisant un stock viral préparé à partir de liquide allantoïdien, un passage supplémentaire de ce stock a été réalisé sur œuf embryonné afin de comparer les populations virales avant et après un passage sur œuf. En dépit de certaines variations, une majorité de variants génétiques présents dans la population virale avant passage sur œuf était retrouvée après passage sur œuf (Figure 14, points rouges).

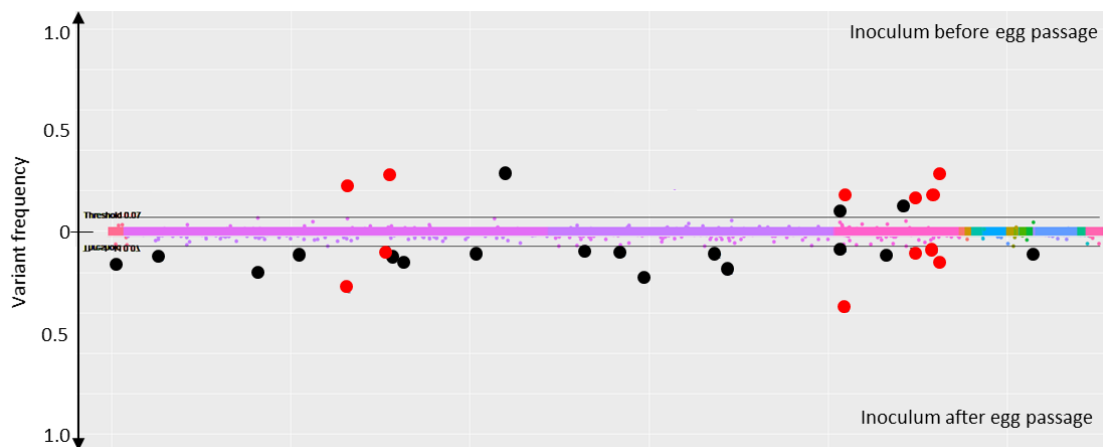


Figure 14: Comparaison des populations virales avant et après un passage sur œuf de poules. En haut, la carte des variants génétiques présents dans la population virale d'IBV avant passage sur œuf ; en bas, la carte des variants génétiques présent dans la population virale après passages sur œuf. Chaque point (noir et rouge) correspond à un variant génétique retrouvé dans la population virale. En rouge sont représentés les variants génétiques commun aux deux population virales. Six variants génétiques sur les neuf détectés avant passage sur œuf, sont également retrouvés après un passage sur œuf, soit 66%.

En conséquence, il a été décidé que le séquençage de broyats de trachées et de reins infectés avec l'IBV se ferait sur des liquides allantoïdiens récoltés à partir d'œufs de poule embryonnés préalablement inoculés à 9 jours d'âge avec les broyats d'organes infectés.

Même si le problème du séquençage d'IBV dans des échantillons tissulaires a pu être résolu, cette méthode d'enrichissement par inoculation dans la cavité allantoïdienne d'œufs embryonnés ne pouvait pas être appliquée pour le TCoV. En effet, pour multiplier ce virus dans des œufs de dindes, il est nécessaire de l'inoculer dans la cavité amniotique d'un œuf embryonné de 19 jours d'âge, afin que le virus puisse gagner les intestins du fœtus et s'y répliquer. Après

quatre jours d'incubation, les intestins sont récupérés dans un tampon salin, broyés et centrifugés afin que le surnageant constitue la nouvelle suspension virale. Mais à la différence du liquide allantoïdien, les intestins sont des tissus et contiennent de très nombreuses cellules eucaryotes. L'étape de broyage entraîne donc la libération d'importantes quantités de matériel génétique cellulaire, ce qui diminue la proportion relative des ARN viraux minoritaires dans l'échantillon. Afin de résoudre ce problème, il a été développé une méthode de préparation différente des suspensions virales afin d'améliorer le séquençage NGS et qui a été testée avec le TCoV coronavirus entérique aviaire.

En partant de la méthode originale d'amplification des coronavirus entériques aviaires sur œuf embryonné, une seule modification de ce protocole a été apportée, remplaçant l'étape de broyage par une étape de vortexage, permettant ainsi de récupérer les particules virales présentes dans la lumière intestinale uniquement. Le but de cette modification est de maintenir au maximum l'intégrité des cellules composant le tissu intestinal et de limiter la libération des acides nucléiques cellulaires dans l'échantillon final. Cette méthode s'est avérée très efficace, puisqu'après quantification de l'ARN cellulaire et viral par RT-PCR quantitative, une réduction de la quantité d'ARNm GAPDH (gène de ménage utilisé pour refléter l'ensemble des ARN cellulaires) a été constatée, sans réduction de la quantité d'ARN viral. Après NGS, la proportion de lectures d'origine virale a augmenté lorsque les échantillons étaient préparés avec la nouvelle méthode (pour des quantités similaires de données NGS produites) et une meilleure profondeur de couverture était obtenue. Ces résultats ont permis de mettre en évidence un enrichissement des échantillons en ARN viraux, facilitant ainsi le séquençage de coronavirus entériques aviaires, dont le TCoV.

Cette deuxième partie du travail a fait l'objet d'une communication courte soumise à *Journal Of Virological Methods*.

Journal of Virological Methods

A modified in-ovo method for isolating Turkey Coronaviruses yields better Next Generation Sequencing results

--Manuscript Draft--

Manuscript Number:	
Article Type:	Short Communication
Keywords:	Coronavirus Enrichment NGS
Corresponding Author:	Alexandre Flageul Agence nationale de sécurité sanitaire de l'alimentation de l'environnement et du travail: Anses PLOUFRAGAN, Bretagne FRANCE
First Author:	Alexandre Flageul
Order of Authors:	Alexandre Flageul Céline Courtillon Chantal Allée Aurélié Leroux Yannick Blanchard Yves Deleforterie Béatrice Grasland Paul Alun Brown
Abstract:	<p>Currently, turkey coronaviruses (TCoV) are isolated from homogenized intestines of experimentally infected embryos to ensure a maximum recovery of virus particles from all components of the intestines. However, the process of homogenization also ensures a release of an important amount of cellular RNAs into the sample that can hinder downstream viral genome sequencing. This is especially the case for next generation sequencing (NGS) that sequences molecules at random. This characteristic means that the heavily abundant cellular RNA in the sample drowns out the minority viral RNA during the sequencing process and consequently very little to no viral genome data is obtained.</p> <p>To address this problem, a method was developed, in which the European strain of TCoV was recovered uniquely from the intestinal lumen without homogenization of the tissue. Based on 87 samples, qRT-PCR showed that compared to the original technique, the new method reduced the mean quantity of cellular RNA without reducing the mean quantity of viral RNA. Importantly, NGS realised on eight of these samples showed that the proportion of viral RNA reads and coverage depth over TCoV genome were increased using the new method.</p>
Suggested Reviewers:	Mark W. Jackwood, Professor mjackwoo@uga.edu working on turkey coronavirus Brian Jordan, Ph.D brian89@uga.edu working on turkey coronavirus



Editor of Journal Of Virological Methods

Objet : short communication submission to Journal of Virological Method

Ploufragan, le 08/10/2020

Laboratoire de Ploufragan-
Plouzané-Niort
Site de Ploufragan

Dear Editor

Unité
Virologie Immunologie
Parasitologie Aviaire et
Cunicoles

We would be grateful if you would accept our short communication entitled '**A modified *in-ovo* method for isolating Turkey Coronaviruses yields better Next Generation Sequencing results**' for consideration in your journal.

Dossier suivi par :
Alexandre Flageul

This paper describes the results of an *in-ovo* method for the isolation of turkey coronavirus that drastically improved downstream next generation sequencing (NGS) of the genome of the isolated virus.

Ligne directe :
+33(0)2 96 01 01 61

The study provides a cost effective very simple method to improve sequencing of turkey coronavirus, and possibly other avian enteric pathogens. The method focused on recovering viral particles from the luminal compartment of avian embryo intestines rather than homogenized embryo intestines as is usually performed. Doing so, the method drastically reduced the amount of cellular RNA, which importantly increased the sequencing of viral RNA.

E- mail :
alexandre.flageul@anses.fr

Zoopôle
BP53
22440 Ploufragan
FRANCE

Yours sincerely,
Alexandre Flageul

Tel: +33(0)2 96 01 62 22
www.anses.fr

REPUBLIQUE FRANÇAISE

Highlights

Highlights

- High-throughput sequencing genome of Turkey coronavirus
- New method of sample preparation to improve NGS
- Viral RNA enrichment method

1 **A modified *in-ovo* method for isolating Turkey Coronaviruses yields better Next**
2 **Generation Sequencing results**

3
4 Alexandre Flageul¹, Céline Courtillon¹, Chantal Allée¹, Aurélie Leroux², Yannick Blanchard²,
5 Yves Deleforterie³, Béatrice Grasland¹ and Paul Alun Brown¹

6
7 ¹Agence National de Sécurité Sanitaire, de l'environnement et du travail (ANSES) laboratory of Ploufragan-
8 Plouzané-Niort, Virology, immunology and parasitology in poultry and rabbit (VIPAC) Unit, Université de
9 Rennes 1.

10 ²Agence National de Sécurité Sanitaire, de l'environnement et du travail (ANSES), laboratory of Ploufragan-
11 Plouzané-Niort, Viral Genetic and Biosafety (GVB) Unit, Université de Rennes 1.

12 ³Hendrix Genetics, les Appats 22150 Plouguenast-Langast
13

14 **Keywords :** Coronavirus, Enrichment, NGS
15
16

17
18 **Abstract**
19

20
21
22 Currently, turkey coronaviruses (TCoV) are isolated from homogenized intestines of
23
24 experimentally infected embryos to ensure a maximum recovery of virus particles from all
25
26 components of the intestines. However, the process of homogenization also ensures a release
27
28 of an important amount of cellular RNAs into the sample that can hinder downstream viral
29
30 genome sequencing. This is especially the case for next generation sequencing (NGS) that
31
32 sequences molecules at random. This characteristic means that the heavily abundant cellular
33
34 RNA in the sample drowns out the minority viral RNA during the sequencing process and
35
36 consequently very little to no viral genome data is obtained.

37 To address this problem, a method was developed, in which the European strain of TCoV was
38
39 recovered uniquely from the intestinal lumen without homogenization of the tissue. Based on
40
41 87 samples, qRT-PCR showed that compared to the original technique, the new method reduced
42
43 the mean quantity of cellular RNA without reducing the mean quantity of viral RNA.
44
45 Importantly, NGS realised on eight of these samples showed that the proportion of viral RNA
46
47 reads and coverage depth over TCoV genome were increased using the new method.
48
49
50

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
34 Turkey Coronaviruses (TCoV) belong to the *Coronaviridae* family, subfamily
35 *Orthocoronavirinae*, genus *Gammacoronavirus*, subgenus *Igacovirus*, and is an important viral
36 agent involved in multifactorial enteritis in turkeys (Guy, 2017). Isolating these viruses has thus
37 far been achieved by inoculation of field samples into the intra amniotic cavity of turkey
38 embryonated eggs (Guionie et al., 2013; Guy, 2015). In this procedure, intestines of
39 embryonated eggs are collected four days post inoculation, homogenized and clarified by
40 centrifugation to produce a virus stock. Whilst this technique means that there is a maximum
41 recovery of infectious virus from the tissue it also means that there is a massive release of
42 cellular nucleic acid into the virus stock that can largely outnumber viral nucleic acids. In terms
43 of “classical virology” (purification, titration, antigen preparation, etc...) this does not pose a
44 problem however, a sample of this type is not the most optimal for next generation sequencing
45 (NGS) of viral genome molecules. This is because NGS is based on random nucleotide
46 sequencing, and thus nucleic acids at the higher ratio in the sample (in this case cellular RNAs)
47 are preferentially sequenced and viral RNA sequencing output is drastically reduced.

25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48 To circumvent this problem, enrichment methods that increase the proportion of a targeted
49 genome can be used. One way is to amplify the targeted genome using Polymerase Chain
50 Reaction (PCR) (Zhou et al., 2009) however, this technique is sensitive to i) PCR efficiency in
51 amplifying the target molecule and to ii) the incorporation of erroneous nucleotides that lead to
52 erroneous downstream sequences. A second way is to remove undesired genetic material (in
53 this case cellular), either by purifying viral particles, using ultracentrifugation (Kuo et al.,
54 2016), gel exclusion size (Loa et al., 2002), or micro-fluidic isolation (Han et al., 2015), or by
55 purifying viral genomes using specific nucleotide primers coupled to magnetic beads (Xing et
56 al., 2016). The major problem with these purification methods is that they all require sufficient
57 amounts of starting sample to overcome an important loss of biological material. This makes
58 them unsuitable when dealing with limited amounts of sample or samples in which the target is
59 at low abundance. Also, all of these methods can be very costly.

47
48
49
50
51
52
53
54
55
56
57
60 In the present study a modification of an existing *in-ovo* method for isolating avian enteric
61 coronaviruses (Guionie et al., 2013) was used in the interest of improving downstream NGS.
62 The modifications of the method focused on removing cellular nucleic acids by eliminating the
63 process of tissue homogenization and on creating an efficient protocol for recovery of viral
64 particles uniquely from the intestinal lumen. Samples prepared using the modified method were
65 compared with samples prepared using the original method.

58
59
60
61
66 To develop the new method, turkey embryonated eggs, provided by the Anses’ department for
67 breeding and experimentation in poultry and rabbits (SELEAC) of the laboratory of Ploufragan-

68 Plouzané-Niort, were inoculated with the European TCoV (strain 080385d) (Guionie et al.,
69 2013). Viral isolate was diluted 1/10 in Phosphate Buffer Saline (PBS) supplemented with
70 penicillin (10U/ μ L), streptomycin (10 μ g/ μ L) and amphotericin B (250 μ g/ μ L). One hundred
71 microliters of TCoV were inoculated into the amniotic cavity of eighty-seven 19-day-old turkey
72 embryonated eggs as described previously (Guionie et al., 2013). After four days of incubation
73 at 37°C, 23-day-old embryos were euthanized and intestines were harvested in 400 μ L of PBS
74 before being vortexed for one minute at maximum speed (2500/min) and centrifuged (3300g,
75 4°C, 15 minutes). Supernatants were harvested and constituted a luminal virus stock (LVS).

76 The pelleted intestines were treated as previously described (Guionie et al., 2013) to produce a
77 tissue virus stock (TVS). Briefly, intestines were transferred into 400 μ L of PBS and frozen at
78 80°C. After thawing, intestines were homogenized using the Tissue Lyser (QIAGEN) for three
79 minutes at 30 shakes per second and centrifuged at 3300g for 15 minutes at 4°C. Supernatants
80 constituted the TVS. These two methods are shown graphically side by side in Figure 1. A total
81 of 87 intestines were harvested from 23-day-old TCoV-infected embryonated turkey eggs,
82 generating 87 LVS and 87 TVS.

83 RNA extractions were performed using the King Fisher automatic extractor. Briefly, 200 μ L of
84 viral stock were lysed with 300 μ L of RLT buffer (QIAGEN) supplemented with 3 μ L of β -
85 mercaptoethanol (SigmaAldrich) and 16 μ g of glycogen (Invitrogen). After lysis, 20 μ L of
86 magnetic beads suspension and 400 μ L of isopropanol were added for RNA precipitation. Three
87 washing steps were realised using washing buffers and ethanol 80%. Finally, RNA was eluted
88 using 100 μ L of elution buffer. Beads, washing and elution buffers were provided in the
89 NucleoMag Vet kit (Macherey-Nagel).

90 TCoV viral RNAs from both LVS and TVS were quantified using a pan avian coronavirus qRT-
91 PCR (Callison et al., 2006). Levels of cellular RNA in the same samples were evaluated using
92 a Power SYBR® Green RNA-to-CT™ 1-Step Kit (Applied Biosystems) targeting the
93 housekeeping gene GAPDH mRNA (Lemaitre et al., 2018). As a normalization method, the
94 same volume of extracted RNA was used in all qRT-PCR so that comparison of both TVS and
95 LVS could be possible. RNA quantification of both type of samples were compared with a
96 paired mode Wilcoxon test from the stat package (v3.5.2) due to the absence of normal
97 distribution of the values, using R (v3.5.2) and R Studio (v1.1.463). P-values below 10⁻³ were
98 considered as significant.

99 After qRT-PCR, it was observed that concentrations of GAPDH mRNA were significantly
100 reduced in LVS compared to TVS, with a mean reduction of 5.07 of the cycle threshold (Ct)
101 (Figure 2A). Concentrations of viral RNA were significantly higher in the LVS compared to

102 TVS, with a mean increase of 0.96 of Ct (Figure 2A). Taken together, these results suggest that
1 103 the modified method indeed removed part of the cellular RNAs of the sample, and that similar
2 104 levels of viral RNA were present in the intestinal lumen and cells of the intestines. Furthermore,
3 105 as GAPDH is only one representative of the complex population of cellular RNAs it could be
4 106 hypothesized that a similar level of reduction occurred for other cellular RNAs molecules. To
5 107 evaluate the capacity of the new method to yield better NGS results, LVS and TVS sequencing
6 108 data obtained from the same intestine (LVS/TVS couple) were compared. In total, four
7 109 randomly chosen LVS/TVS couples were compared. First, total RNA concentration was
8 110 determined for all chosen samples using fluorimeter Qubit 2.0 device (Invitrogen) and the
9 111 RNA-HS Assay kit (Invitrogen). It was observed that total RNA concentrations were markedly
10 112 lower in the LVS (between 24.5 and 48.8 ng/ μ L) than in TVS (between 318 and 853 ng/ μ L),
11 113 except for one LVS which had a RNA concentration (20.3 ng/ μ L) similar to that of its coupled
12 114 TVS (20.0 ng/ μ L) (Figure 2B). This observation was further evidence that the new method,
13 115 avoiding tissue homogenization, had substantially reduced the release of cellular RNAs into the
14 116 sample.

117 To perform NGS and prepare RNA libraries, a maximum concentration of RNA was used
118 respecting an upper ceiling of 500ng in a maximum volume of 20 μ L and was treated with four
119 units of DNase (Turbo DNA-*free*TM kit from Life Technologies). Following this treatment,
120 ribosomal RNAs were removed (ribodepletion) using Low Input RiboMinusTM Eukaryote
121 System v2 (Life Technologies). RNA libraries were then prepared with Ion Total RNA-seq kit
122 2 (Life Technologies). Briefly, this process involved enzymatic fragmentation of RNA and
123 ligation of adaptors to the fragmented RNA to generate cDNA from RNA molecules by reverse
124 transcription. Resulting cDNAs were then amplified and barcoded for identification. Finally,
125 library quality was checked on a Bioanalyzer from Agilent (High Sensitivity DNA Kit and RNA
126 6000 Pico Kit, Agilent) and submitted to Ion-Torrent sequencing. For the bioinformatic process,
127 read trimming was realised with Trimmomatic (v0.38) using default parameters (Bolger et al.,
128 2014). The genomic composition of the samples was determined by aligning clean NGS data
129 on the NCBI nt database (last update, November 2018) using Megablast (2.2.26) with default
130 parameters (Morgulis et al., 2008) and visualised with KronaTools (v2.7) (Ondov et al., 2011).
131 Additionally, reads were aligned on the previously published genome of TCoV-080385d
132 (GenBank accession number KR822424.1, Brown et al., 2016) using BWA (v0.7.17-r1188)
133 and Samtools (v1.9.9).

134 The total number of produced reads were between 1282650 and 3507514, depending on the
135 sample. For all LVS/TVS couples, the proportion of viral reads was higher in LVS (Figure 2C),

136 with a mean proportion of viral reads in the LVS of 83.25% (+/- standard deviation (SD) 3.7%)
137 of the total produced reads against 8.07% (+/- SD 12.05%) in the TVS. Other reads were mostly
138 from eukaryotic origin but traces of bacterial reads were also detected (less than 0.3%, data not
139 shown). After aligning reads over the reference genome, coverage depth was evaluated for both
140 LVS and TVS. On average, more total reads (viral and other) were produced for the TVS, with
141 the exception of sample 4, for which there was more reads in the LVS. Despite this, the viral
142 coverage depths were higher for all LVS (light grey traces) than for all the TVS (black traces),
143 with median coverage depths of 455, 929, 180, 180 for LVS (sample 1 to 4 respectively),
144 compared to 1, 23, 0, 4 for TVS (samples 1 to 4 respectively) (Figure 2D).

145 The random sequencing approach of NGS allows sequencing of all nucleic acid molecules
146 present within a sample. However, in practice this random technique of sequencing can make
147 the characterization of minority nucleic acid chains difficult since most of the sequenced data
148 would be determined from the most represented nucleic acid chains. Thus, a complete picture
149 of the minority viral genomic nucleic acid composition of a sample of interest is seldom
150 obtained. In this study, a modified version of the existing Guionie's method (Guionie et al.,
151 2013) successfully increased the proportion of viral RNA compare to RNAs from other origin
152 (eukaryote and bacteria) that ultimately improve NGS results.

153 The present method gives one answer to the following optimisation problem: how to obtain the
154 maximum amount of viral data to sequence correctly a viral genome, while producing the
155 minimum amount of total NGS data. Therefore, this new method is a much-needed *in-ovo*
156 alternative for avian enteric coronavirus propagation and downstream NGS. Compared to other
157 purifying or amplification techniques, this simple and cost effective method does not require
158 important amounts of starting sample, expensive and specific equipment or reactivities and
159 represent a simple adaptation of a well practice and classic virological method in the interest of
160 addressing modern molecular virology questions that require robust sequencing data.
161 Moreover, the fact that this method does not employ any measures that create bias in the
162 sequencing data, such as PCR amplification or genome capture, then sequencing data should
163 be a better representation of the actual genome population.

164 Finally, these results highlight the fact that if viruses replicating *in-vivo* can be collected from
165 cellular secretions rather than the cells themselves then nonspecific NGS sequencing of their
166 genomes should yield better results.

167

168

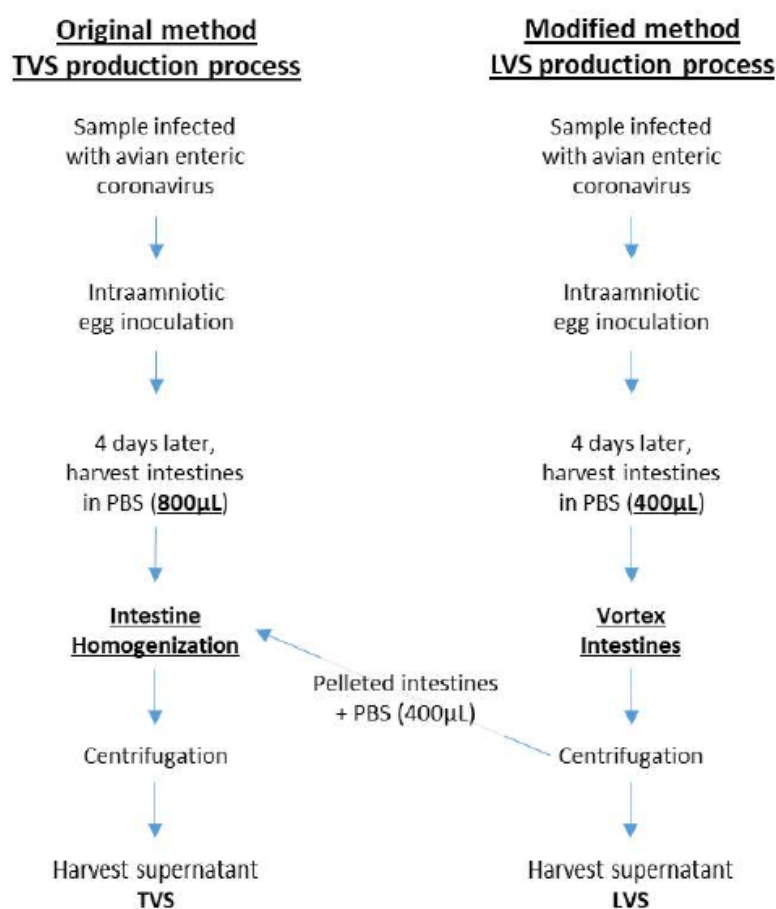
169

170 **Acknowledgements**

171 This work was partly funded by Agence National de Sécurité Sanitaire, de l'Environnement et
172 du travail (ANSES), the "Région Bretagne", the "Université Bretagne Loire" and the
173 "Université de Rennes" (UNIR) (code of convention : 180019) and partly funded by the
174 European Joint Programme One Health METASTAVA (no 773830).

176 **Declaration of interest**

177 None.



179 **Figure 1:** Ideogram of the original and modified method of TCoV
180 isolation. TVS = tissue virus stock, LVS = lumen virus stock, **bold and**
underlined text = differences.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61

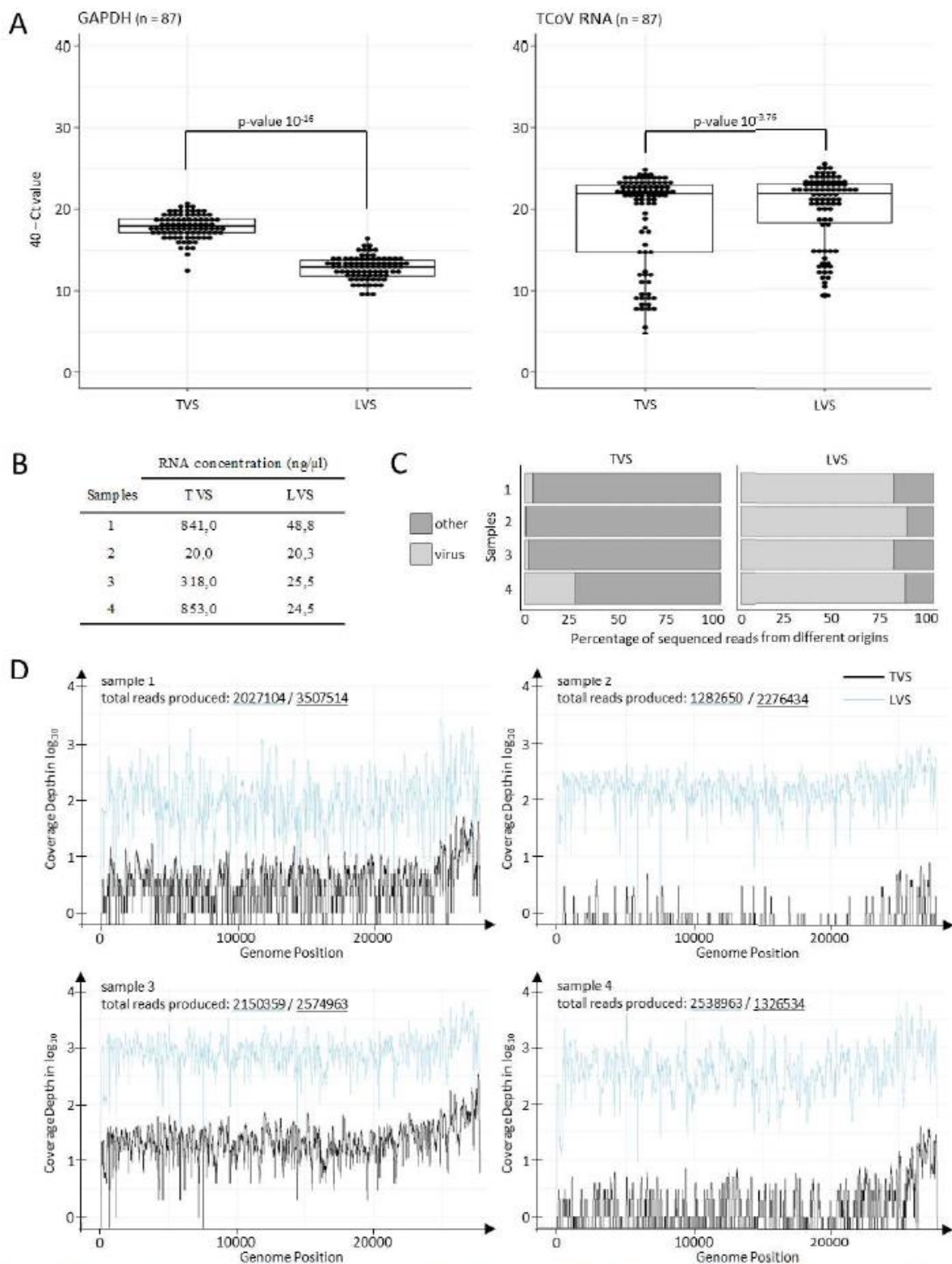


Figure 2: Quantification and NGS read ratio of cellular and TCoV viral RNA. A) Comparison of the concentration of GAPDH mRNA and TCoV RNA. Y-axis represents reversed Ct values (40 - Ct), meaning that lower values equal lower concentration. Statistical tests were performed using unparametrical Wilcoxon test in paired mode. B) Total RNA concentrations of the sequenced samples. C) Relative proportion of viral/other (cellular and bacterial) reads of four randomly chosen TVS/LTS couples that were sequenced. D) Coverage depths obtain for both TVS (black traces) and LTS (light grey traces) after aligning NGS data to the TCoV-080385d reference genome (GenBank accession number KR822424.1). Results were formatted and visualised using the R package ggplot2 (v3.1.0).

182 **References**

- 183 Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data.
184 *Bioinformatics* 30, 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
- 185 Brown, P.A., Touzain, F., Briand, F.X., Gouilh, A.M., Courtillon, C., Allée, C., Lemaitre, E., De Boissésou, C.,
186 Blanchard, Y., Eterradosi, N., 2016. First complete genome sequence of European turkey coronavirus
187 suggests complex recombination history related with US turkey and guinea fowl coronaviruses. *J. Gen.
188 Virol.* 97, 110–120. <https://doi.org/10.1099/jgv.0.000338>
- 189 Callison, S.A., Hilt, D.A., Boynton, T.O., Sample, B.F., Robison, R., Swayne, D.E., Jackwood, M.W., 2006.
190 Development and evaluation of a real-time Taqman RT-PCR assay for the detection of infectious
191 bronchitis virus from infected chickens. *J. Virol. Methods* 138, 60–65.
192 <https://doi.org/10.1016/j.jviromet.2006.07.018>
- 193 Guionie, O., Courtillon, C., Allee, C., Maurel, S., Queguiner, M., Eterradosi, N., 2013. An experimental study
194 of the survival of turkey coronavirus at room temperature and +4°C. *Avian Pathol.* 42, 248–252.
195 <https://doi.org/10.1080/03079457.2013.779364>
- 196 Guy, J.S., 2017. Viral Enteric Infections, in: *Diseases of Poultry*, 13th Edition. John Wiley & Sons, Ltd, pp. 375–
197 415. <https://doi.org/10.1002/9781119421481.ch12>
- 198 Guy, J.S., 2015. Isolation and Propagation of Coronaviruses in Embryonated Eggs, in: Maier, H.J., Bickerton, E.,
199 Britton, P. (Eds.), *Coronaviruses: Methods and Protocols*, *Methods in Molecular Biology*. Springer,
200 New York, NY, pp. 63–71. https://doi.org/10.1007/978-1-4939-2438-7_7
- 201 Han, H.-S., Cantalupo, P.G., Rotem, A., Cockrell, S.K., Carbonnau, M., Pipas, J.M., Weitz, D.A., 2015. Whole-
202 Genome Sequencing of a Single Viral Species from a Highly Heterogeneous Sample. *Angewandte
203 Chemie International Edition* 54, 13985–13988. <https://doi.org/10.1002/anie.201507047>
- 204 Kuo, L., Hurst-Hess, K.R., Koetzner, C.A., Masters, P.S., 2016. Analyses of Coronavirus Assembly Interactions
205 with Interspecies Membrane and Nucleocapsid Protein Chimeras. *Journal of Virology* 90, 4357–4368.
206 <https://doi.org/10.1128/JVI.03212-15>
- 207 Lemaitre, E., Allée, C., Vabret, A., Eterradosi, N., Brown, P., 2018. Single reaction, real time RT-PCR detection
208 of all known avian and human metapneumoviruses. *J Virol Methods* 251, 61–68.
209 <https://doi.org/10.1016/j.jviromet.2017.10.010>
- 210 Loa, C.C., Lin, T.L., Wu, C.C., Bryan, T.A., Thacker, H.L., Hooper, T., Schrader, D., 2002. Purification of
211 turkey coronavirus by Sephacryl size-exclusion chromatography. *J. Virol. Methods* 104, 187–194.
212 [https://doi.org/10.1016/s0166-0934\(02\)00069-1](https://doi.org/10.1016/s0166-0934(02)00069-1)
- 213 Morgulis, A., Coulouris, G., Raytselis, Y., Madden, T.L., Agarwala, R., Schäffer, A.A., 2008. Database indexing
214 for production MegaBLAST searches. *Bioinformatics* 24, 1757–1764.
215 <https://doi.org/10.1093/bioinformatics/btn322>
- 216 Ondov, B.D., Bergman, N.H., Phillippy, A.M., 2011. Interactive metagenomic visualization in a Web browser.
217 *BMC Bioinformatics* 12, 385. <https://doi.org/10.1186/1471-2105-12-385>
- 218 Xing, N., Guan, X., An, B., Cui, B., Wang, Z., Wang, X., Zhang, X., Du, Q., Zhao, X., Huang, Y., Tong, D.,
219 2016. Ultrasensitive Detection of Porcine Epidemic Diarrhea Virus from Fecal Samples Using
220 Functionalized Nanoparticles. *PLoS One* 11. <https://doi.org/10.1371/journal.pone.0167325>
- 221 Zhou, B., Donnelly, M.E., Scholes, D.T., St. George, K., Hatta, M., Kawaoka, Y., Wentworth, D.E., 2009.
222 Single-Reaction Genomic Amplification Accelerates Sequencing and Vaccine Production for Classical
223 and Swine Origin Human Influenza A Viruses. *J Virol* 83, 10309–10313.
224 <https://doi.org/10.1128/JVI.01109-09>
- 225

Authorship statements

Alexandre Flageul: Investigation, Formal analysis, Validation, Visualization, Writing - Original Draft

Céline Courtillon: investigation

Chantal Allée: investigation

Aurélié Leroux: resource

Yannick Blanchard: resource

Yves Deleforterie: resource

Béatrice Grasland: Supervision, Writing - Review & Editing, Project administration, Funding Acquisition

Paul Alun Brown: Supervision, Methodology, Writing - Review & Editing, Project administration, Funding Acquisition

8 Evolution des *Gammacoronavirus* aviaires TCoV et IBV en absence et en présence d'immunité vaccinale préexistante

Après avoir mis en place les méthodes et outils nécessaires au séquençage et à la détection de variants génétiques dans les populations virales, les analyses évolutives des coronavirus aviaires TCoV et IBV ont pu être conduites.

À travers ces analyses, deux questions ont été posées :

- i) « **Comment évoluent les *Gammacoronavirus* aviaires au cours de plusieurs passages sur leurs hôtes ?** » Pour répondre à cette question, l'évolution du TCoV a été suivie au cours de dix passages sur dindes en absence d'immunité pre-existante, et l'évolution de l'IBV a été suivi au cours de trois passages sur poulets en absence d'immunité pré-existante
- « **L'immunité préexistante vaccino-induite modifie-t-elle l'évolution des *Gammacoronavirus* aviaires ?** » Pour répondre à cette question, l'évolution de l'IBV a été suivie au cours de trois passages sur poulets en présence d'une immunité pré-existante vaccino-induite et comparer à l'évolution sur poulets sans immunité pré-existante.

8.1 Evolution du TCoV en absence d'immunité vaccinale pré-existante

Au cours d'une étude précédente, il a été montré que le Fr-TCoV était hautement transmissible et la dose minimale infectante est extrêmement faible, inférieure aux limites de détections des méthodes moléculaires (Brown et al., 2018). Il a également été montré que des animaux infectés par le Fr-TCoV excrétaient du virus infectieux pendant une période d'au moins six semaines après l'infection (Brown et al., 2018). Au vu des caractéristiques du TCoV et en l'absence de mesure de contrôle de l'infection par TCoV comme la vaccination, il était important de caractériser l'évolution du TCoV afin de savoir si le TCoV pouvait générer rapidement diverses populations de génomes et de nouveaux variants génétiques. Avec cet objectif, dans cette étude, l'évolution de la souche Fr-TCoV (stock viral issu de six passages sur œufs de dinde) a été caractérisée par NGS au cours de dix passages du virus sur dindes EOPS.

Les expérimentations sur animaux ont été réalisées dans les installations de biosécurité de niveau 2 du laboratoire ANSES de Ploufragan-Plouzané-Niort, conformément aux réglementations européennes et françaises sur le bien-être animal. Les dindes EOPS provenaient du service d'élevage et d'expérimentation avicole du laboratoire ANSES de Ploufragan-Plouzané-Niort. Le protocole expérimental décrit ci-dessous a été approuvé (Référence APAFIS # 10839-

2017073115311452 v4)) par le comité d'éthique Anses / ENVA / UPEC enregistré sous le numéro 16 et par le ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation.

Un premier lot de quinze dindes EOPS âgées de 10 jours d'âge (j.a.) a été écouvillonné par le cloaque afin de déterminer l'absence d'infection à TCoV. Puis, trois dindes de 10 j.a. ont été choisies au hasard (nommées A1, B1 et C1) et individuellement placées dans trois enclos séparés par des bâches en plastique au sein de l'animalerie 1. Le même jour, les dindes A1, B1 et C1 ont été inoculées par voie orale avec 250 µL (250 DIE50) de Fr-TCoV (nommé ensuite inoculum TCoV). Deux jours plus tard, des écouvillons cloacaux ont été prélevés sur les sujets inoculés et les écouvillons ont été remis en suspension dans 1 mL de milieu de culture MEMH. Ces échantillons ont été nommés A1_CS (écouvillon cloacal A1), B1_CS et C1_CS. Après écouvillonnage, les trois dindes (A1, B1 et C1) ont été euthanasiées par électronarcose suivie d'une saignée et les intestins ont été prélevés et conservés à -80°C.

De nouveau, trois animaux choisis au hasard (A2, B2 et C2) ont été déplacés dans les mêmes enclos de l'animalerie 1 et ont été inoculés par voie orale avec 250 µL de A1_CS, B1_CS et C1_CS respectivement. Deux jours plus tard, des écouvillons cloacaux ont été prélevés sur les sujets inoculés afin de préparer les inocula du troisième passage (A2_CS, B2_CS et C2_CS) de la même façon que précédemment. Puis les dindes ont été euthanasiées afin de prélever les intestins. Le troisième passage a été réalisé dans une nouvelle animalerie (animalerie 2) contenant trois enclos séparés par des bâches en plastique et dans lesquels les dindes A3, B3 et C3 ont été individuellement placées et inoculées avec A2_CS, B2_CS et C2_CS. Deux jours plus tard, l'ensemble du protocole a été répété. Ceci a été effectué jusqu'au passage 5, à partir duquel un nouveau lot de quinze dindes EOPS de 10 j.a. (également négatives pour la présence de TCoV avant inoculation) a été utilisé pour réaliser cinq passages supplémentaires. Au total, dix passages ont été réalisés en suivant le même protocole. Tous les deux passages, chaque animalerie a été nettoyée et désinfectée à l'aide de peroxyde d'hydrogène et d'acide peracétique.

Au cours de cette expérimentation, une excrétion virale a été détectée dans les écouvillons cloacaux de tous les passages. Aucun signe clinique n'a été observé chez les dindes à aucun des passages.

Dans le but de procéder à des analyses de variants génétiques, les contenus des caeca issus des dindes utilisées lors des passages 1, 5 et 10 ont été passés une fois sur œuf de dinde embryonné et les intestins des embryons ont été récupérés et préparés avec la méthode nouvellement mise au point afin de séquencer plus facilement les populations virales en NGS. Après traitement des données NGS avec le pipeline nouvellement développé, cinq modifications (notées Cch1, 2, 3, 4

et 5, Tableau 2) de la séquence consensus de l'inoculum TCoV sont apparues dès le premier passage sur dindes et ont également été détectées après les passages 5 et 10 dans les trois répliques, indiquant une conservation de ces mutations tout au long de l'expérience. D'autres variations génétiques ou changements de consensus ont été observés lors des différents passages, mais aucune autre variation n'a été retrouvée lors des trois passages séquencés. Cette rapidité de sélection de variants pourrait être le résultat d'une sélection de variants génétiques présents dans la population virale de l'inoculum TCoV. Cependant, il n'est pas possible de confirmer ou d'infirmer cette hypothèse car la population virale de l'inoculum TCoV n'a pu être correctement déterminée du fait d'une quantité insuffisante de données NGS.

Tableau 2 : Variants génétiques identifiés après sélection au cours des passages sur dindes lors de l'expérimentation Fr-TCoV

	Position génomique	Nucléotides consensus de l'inoculum	Nucléotide(s) sélectionné(s) après passage	Région protéique	Impact protéique
Cch1	9901	C	T	PP1ab – NSP6	H-3125-Y
Cch2	10450	T	C	PP1ab – NSP6	F-3308-L
Cch3	15349	A	C	PP1ab – NSP13	Q-4941-T
Cch4	20951	G	T	S1	S-196-I
Cch5	21737	G	A	S1	S-458-Q

À partir d'un unique stock de virus issu de six passages sur œufs, les mêmes modifications conservées de séquences consensus ont été retrouvées dans les trois répliques. Un tel résultat pourrait indiquer que l'évolution naturelle d'un coronavirus aviaire n'est pas aléatoire mais qu'elle tendrait vers un génome optimal adapté à la réplication dans l'hôte. Cela signifierait qu'à partir d'une population virale donnée, quel que soit le nombre d'inoculations effectuées, l'évolution du virus serait la même tant que les conditions de réplication restent inchangées. Cependant, comme les répétitions de chaque passage ont été effectuées dans la même animalerie, mais dans des enclos différents séparés par des bâches en plastique, il ne peut pas être exclu que des virus excrétés par les animaux aient pu être transmis par voie aérienne d'un enclos à un autre. Mais cela semble peu probable car des travaux antérieurs ont conclu que la transmission aérienne du TCoV n'était pas possible au-delà de 2-3 mètres de distance (Brown et al., 2018), indiquant donc que les bâches en plastique devaient être suffisantes pour éliminer le risque de transmission aérienne.

L'impact des mutations sur la fonctionnalité des protéines a été analysé. Cch1, Cch2 et Cch3 ont provoqué des changements de séquences protéiques de NSP6, NSP6 et NSP13 respectivement, mais il semble que ces changements ne modifient pas profondément la structure de la protéine. En effet, les séquences variantes de Cch1 et Cch2 sont retrouvées dans des

séquences IBV de la base de données du NCBI, et la séquence variante de Cch3 ne modifie pas la polarité ni la charge de la protéine à cette position. Cch4 et Cch5 ont provoqués une modification de la séquence de la protéine de spicule au niveau du domaine S1. Cch4 a remplacé une sérine (acide aminé polaire non chargée) par une isoleucine (hydrophobe). Cette position protéique est localisée à la surface de la protéine de spicule, et comme ces deux acides aminés ont des propriétés biochimiques différentes, la modification pourrait potentiellement avoir un impact important sur la structure de la protéine et, par exemple, pourrait induire un changement de la fixation de la protéine à son récepteur ou de la liaison à des anticorps neutralisants. Cch5 a remplacé une sérine par une asparagine, deux acides aminés polaires non chargés. Cette modification est donc supposée ne pas modifier considérablement la conformation de la protéine. Ainsi, en se basant uniquement sur la fonctionnalité des protéines, les raisons pour lesquelles ces mutations ont été sélectionnées dans tous les répliques restent inconnues.

En conclusion, ces travaux ont montré que cinq positions dans le génome viral ont été modifiées au niveau du consensus dès le premier passage et ont été fixées jusqu'au dixième passage. Par ailleurs, les changements de séquences consensus maintenu au cours des passages ont été observés au cours de trois répliques, indiquant que l'évolution du TCoV n'est pas aléatoire.

Ce travail sera repris pour une valorisation par un article qui sera soumis à *Transboundary and Emerging Diseases*.

Evolution of Turkey Coronavirus genome populations over ten passages in specific pathogen free turkeys.

Introduction

Turkey coronavirus (TCoV) is a *Gammacoronavirus* that belongs to the newly established *Igacovirus* subgenus (ICTV). TCoV has been described to be the principal viral agent of a multifactorial enteric disease of turkeys which is known as “poult enteritis complex” (PEC) (Barnes et al., 2000; Gomaa et al., 2008). PEC is found everywhere where turkeys are farmed and occurs in all ages of birds however, younger birds seem to be most susceptible. In certain cases this disease can be severe resulting in high mortality and in these cases is given the name “poult enteritis mortality syndrome” (PEMS). The most striking example of PEMS occurred in the United States of America between 1991 and 1997 and cost the industry an estimated 160 million dollars (Barnes et al., 2000).

TCoV, much like its closest relative infectious bronchitis virus (IBV) with which it shares all but the spike glycoprotein S gene of its genome, is extremely infectious. Under experimental conditions, TCoV has been shown to be transmitted from one infected turkey to another every 2.5h (Brown et al., 2018) and to be infectious at a dose beyond the limits of detection by real-time RT-PCR (Brown et al., 2018). In addition to these favourable characteristics for infection, the virus is also excreted as infectious virions for at least six weeks (Brown et al., 2018). In practical terms these characteristics highlight why TCoV infections are difficult to control once an infectious episode is launched and in virological terms highlight a virus that has great opportunity to generate and sustain diverse genome populations i) through its rapid infectious rate and ii) through the duration of infection. To date, the potential for TCoV to quickly generate diverse genome populations and new genetic variants remains a hypothesis, as limited full-length genome data exists in the database to correctly assess TCoV diversity in the field and as no experiments have been designed to test it under controlled conditions. However, field data on IBV genomes and recent experimental findings for IBV (Flageul et al., submitted-a), which showed rapid emergence and selection of genetic IBV variants upon passage in specific pathogen free (SPF) chickens under experimental conditions would seem to indirectly support the hypothesis.

The current study was thus designed to assess the evolution of TCoV genome populations during passages in SFP turkeys. Specifically, the European isolate of TCoV (Fr-TCoV) was inoculated via the oral route to three turkey poults. Then, ten passages were performed for each of the three initial birds. The evolution of full-length viral genomes replicating in caecal tissues of the three birds from passage 1, the three from passage 5 and the three from passage 10 was determined by

high-throughput sequencing and processing data with VVV, the newly developed bioinformatics tool designed to detect viral genetic variants (Flageul et al., 2021).

Material and methods

Virus

Fr-TCoV 080385d/6.3, was titrated on eggs [$10^{4.01}$ egg infectious dose (EID) 50 / mL] during a previous experiment (Brown et al., 2018). To constitute the inoculum of the first passage this virus was diluted in MEMH [MEM 9,66 g/L ; penicillin 50000 U/L ; streptomycin 0.05 g/L ; hepes 20 mM ; NaHCO₃ 8.85 g/L; diluted in ultrapure water, pH 7.2 adjusted with NaOH 1N] to adjust the viral titer to 10^3 EID50/mL and was next referred as TCoV-inoculum.

Experimental trial

Animal experiments were performed in accordance with European Union and French regulations on animal welfare in experiments. The protocol was approved (Reference APAFIS#10839-2017073115311452 v4) by the ANSES/ENVA/UPEC ethical committee registered under number #16. Experiments were carried out in ANSES' biosafety level 2 facilities. Specific pathogen free (SPF) turkeys were provided by the Anses' department for breeding and experimentation in poultry of the laboratory of Ploufragan-Plouzané-Niort.

One batch of fifteen 10 day-old SPF turkeys were used. Cloacal swabs were taken to check the absence of avian enteric coronavirus infection at the start of the experiment.

At the day of the inoculation (day 0), three randomly chosen 10 day-old turkeys (named A1, B1 and C1) were placed in room 1 in three individual side by side pens separated at a height of one meter by plastic tarpaulin. A1, B1 and C1 were then orally inoculated with 250 EID50/bird of TCoV-inoculum. Cloacal swabs were collected two days post-inoculation on each bird and resuspended in 1 ml of MEMH. These samples were named A1_CS (A1 cloacal swab), B1_CS and C1_CS. The three turkeys (A1, B1 and C1) were euthanized by electronarcosis followed by bleeding two days post-inoculation. Intestines were sampled and stored at -80°C .

Then, three randomly chosen animals (A2, B2 and C2) were moved in the same pens of room 1 and were inoculated orally using 250 μL of A1_CS, B1_CS and C1_CS respectively. Cloacal swabs were collected two days post-inoculation on these three birds to prepare A2_CS, B2_CS and C2_CS inocula for the third passage as previously described. Animals were euthanized and intestines were sampled.

The third passage was carried out in a new clean room 2 with three separate pens with turkeys A3, B3 and C3 that were inoculated with A2_CS, B2_CS and C2_CS. Cloacal swabs were collected, A3_CS, B3_CS and C3_CS were prepared and the same process as earlier was applied. This was

performed until passage 5. At that time, a new batch of fifteen 10-day old turkeys were used for the five further passages. Ten passages were performed in total using the same protocol.

Every two passages, each room was cleaned and disinfected using H₂O₂ and acid peracetic.

Preparation of intestinal contents for analysis of the variant population

For each intestines, contents from the caecum intestinal segment were individually collected and homogenised in PBS at the concentration of 1 g/mL. The homogenates were vortexed for one minute (min) at maximum speed (2500 shakes/min) and then centrifuged at 3000 g for 15 min at 4°C. Supernatants constituted the clarified intestinal contents and were frozen at -80°C.

Amplification of the viral population in clarified intestinal content by egg inoculation

TCoV-inoculum and clarified intestinal contents were diluted (1/10) in PBS supplemented with penicillin (10 U/μL), streptomycin (10 μg/μL) and amphotericin B (250 μg/μL). After one hour of incubation at 4°C, diluted samples were inoculated in the amniotic cavity of 19 day-old embryonated SPF turkey eggs as previously described (Guionie et al., 2013) Eggs were incubated for 6 days at 37°C and candled every day. Turkey embryos that died during the first 24 hours of incubation were not kept. At the end of the incubation period, or after the death of embryos, intestines were harvested and suspended in 400 μL of PBS before being vortexed for one minute at maximum speed (2500 shakes/min) and centrifuged for 10 min at room temperature (Flageul et al., submitted-b). Supernatants constituted the amplification of individual turkey intestinal contents and were individually frozen at -80°C before RNA extraction.

RNA extraction

For each amplification of individual turkey intestinal contents, RNAs were extracted using the King Fisher automatic extractor using NucleoMag Vet kit (Macherey-Nagel). Briefly, 200 μL of each amplification were lysed with 300 μL of RLT buffer (QIAgen) supplemented with 3 μL of β-mercaptoethanol (SigmaAldrich) and 16 μg of glycogen (Invitrogen). After lysis, 20 μL of magnetic beads and 400 μL of isopropanol were added for RNA precipitation. Three washing steps were realised (two using washing buffers and one using 80 % ethanol). Finally, RNA was eluted using 100 μL of elution buffer.

Quantification of TCoV genome copies by reverse transcription polymerase chain reaction (qRT-PCR)

Quantification of viral RNA was performed using a pan avian coronavirus qRT-PCR as described previously (Callison et al., 2006).

Illumina sequencing and NGS data processing

Samples from passages 2, 3, 4, 6, 7, 8, 9 were not sequenced. Only the samples from TCoV-inoculum, passage 1, 5 and 10 were subjected to sequencing. RNA libraries were prepared using the NEBNext Ultra II Directional RNA library kit. Briefly, ribosomal RNAs were removed with RNaseH and remaining RNAs were converted into double strand DNA (dsDNA) by random priming reverse transcription and second strand DNA synthesis. Then, sequencing adaptors were ligated to dsDNA and a short PCR enrichment was performed before sequencing. NGS was performed using NextSeq Illumina technology.

To determine the consensus sequence of the TCoV-inoculum, NGS data were processed through the Viral Variant Visualiser (VVV) and compared to the previously published sequence of Fr-TCoV (GenBank accession number KR822424.1) (Flageul et al., 2021). TCoV-inoculum consensus genome was deduced from VVV results.

NGS data from TCoV-inoculum and from amplifications of individual turkey intestinal contents were processed twice with VVV. The first analysis involved all the reads in the NGS file while the second analysis removed reads considered as duplicates (identical reads). For reasons previously defined, (Flageul et al., 2021) only genetic variants detected in both analysis and present in the viral population above a threshold value of 7% were analysed. Data were considered robust when the median coverage depth was greater than 500 and when a minimum of 200 reads was present in locations where variants were detected. Other data were annotated as “low coverage”.

Comparison of protein sequences with databases

Predicted protein sequences from observed genomic mutations were compared to the non-redundant database of NCBI using blastp with default parameters. For each of the predicted protein sequence, the 500 better scored matching sequences were used to performed a multiple sequence alignment using the MSA tool of the NCBI.

Results

Clinical symptoms and viral shedding

No viral RNA was detected in the cloacal swabs taken from all animals three days before the first inoculation. No clinical signs (prostration or diarrhoea) were observed in all birds throughout the trial.

The viral shedding was evaluated through the number of genome copies assessed in cloacal swabs (Table 1). The viral genomic load varied between $10^{2.08}$ (one bird at the third passage) and $10^{5.37}$ genome copies/ μ L (one bird at the second passage). However, most of the samples (26/30) presented viral loads ranging between $10^{3.5}$ and 10^5 copies/ μ L.

Viral population in the TCoV-inoculum

Twenty-three genetic variants were detected in both analyses (with and without removing the duplicate reads), but all of them were annotated as ‘low-coverages’, supported by coverage depths below 271 for each of the position where variants were detected (data not shown). Thus, they were not considered as variants using the defined criteria.

Evolution of TCoV in turkey through 10 passages

To assess the evolution of TCoV-inoculum over passages in turkeys, the viral populations in amplifications of individual turkey intestinal contents were determined and compared to each other, for animals A1, A5, A10; B1, B5, B10 and C1, C5, C10 only.

All detected variants were compiled into Figure 1. Each variant is indicated by a blue square, which is light if the variant frequency was less than 50% of the total sequences at that position or which is dark if the variant was above 50% of the total sequences meaning a change of the consensus sequence represented by TCoV-inoculum.

For all samples, median coverage depths were comprised between 4118 and 32849 for the analysis that involved all reads, and between 312 and 417 in the analysis that removed reads considered as duplicates. For all samples, low coverage depth (< 200 reads per genome position) was detected at the 5’ extremity of the genome from position 1 to 53 for the analysis that involved all reads, and from position 1 to 72 for the analysis that removed reads considered as duplicates. Two other regions had low coverage between genome positions 15527 and 15757, and at the 3’ extremity from position 27599 to the end, but no genetic variants were detected in those regions (Figure 1). A total of 49 genetic variants were detected at passages 1, 5 and 10 in replicates A, B and C. In the three replicates, five positions led to a consensus change (noted Cch1, Cch2, Cch3, Cch4, and Cch5) and were selected after passages (Figure 1). Cch1 (modification C-9901-T) and Cch2 (T-10450-C) were detected at the consensus level at passage 1 in replicates A and B and in all replicates at passage 5 and 10. Cch3 (A-15349-C) and Cch4 (G-20951-T) were detected at passage 1 below or at the consensus level in replicate A and B respectively but was not detected in replicate C. However, at passage 5 and 10, Cch3 and Cch4 were detected in all replicates at the consensus level. Cch5 (G-21737-A) was detected at passages 1, 5 and 10 in all replicates at the consensus level.

Predicted protein impact of the non-synonymous selected mutations

Potential impact of the main observed mutations on protein sequence was analysed and modified protein sequences were compared with protein sequences in the NCBI non-redundant database.

The genomic mutation Cch1 (C-9901-T) was predicted to modify the coded protein sequence, replacing a Histidine by a Tyrosine at position 3125 of the polyprotein 1ab (NSP6 region). After blastp and multiple sequence alignment, Histidine and Tyrosine were both found at this position and no other amino acid was detected.

The genomic mutation Cch2 (T-10450-C) was predicted to modify the coded protein sequence, replacing a Phenylalanine by a Leucine at position 3308 of the polyprotein 1ab (NSP6 region). After blastp and multiple sequences alignment, Leucine was the most present amino acid at this position and Phenylalanine was only found in two IBV sequences of the alignment [accession numbers from NCBI protein database: AFJ11174, ALH21109].

The genomic mutation Cch3 (A-15349-C) was predicted to modify the coded protein sequence, replacing an Asparagine by a Threonine at position 4941 of the polyprotein 1ab (NSP13 region). After blastp and multiple sequence alignment, Asparagine was the most present amino acid at this position but Serine was also detected. No Threonine was observed at this position.

The genomic mutation Cch4 (G-20951-T) was predicted to modify the coded protein sequence, replacing a Serine by an Isoleucine at position 196 of the spike protein. After blastp and multiple sequence alignment, Serine was present in Fr-TCoV related sequences [ADK88938, CBA13338, CBB12677] but Glutamic acid was the most present amino acid at this position. Isoleucine was not detected at this position.

The genomic mutation Cch5 (G-21737-A) was predicted to modify the coded protein sequence, replacing a Serine by an Asparagine at position 458 of the Spike protein. After blastp and multiple sequence alignment, Serine was the most present amino acid at this position, but Asparagine was also detected in the three Fr-TCoV related sequences and one duck coronavirus sequence [YP_009825008.1].

Discussion

The present study focused on studying the evolution of the Fr-TCoV over ten passages in turkeys in the absence of selection pressure.

During this experiment, it was observed that modifications of the consensus sequence of TCoV-inoculum appeared in five positions as early as the first passage in turkeys and were not modified after further passages. This rapidity could be the result of a selection of genetic variants present in the viral population of the TCoV-inoculum. However as TCoV-inoculum has not been sequenced

with the minimum required amount of data, the composition of the viral population was not determined with sufficient confidence to validate this hypothesis. Indeed, coverages of the TCoV-inoculum genome ranged from 52 to 165 at positions of Cch1 to Cch5. Therefore, genetic variants might have been present at these positions of the viral genome but were left undetected due to low coverages.

The rapidity of selection echoed previously published work on avian coronaviruses. IBV was passaged three times in chickens and five selected mutations that have been identified to modify the protein sequences appeared as soon as the first passage (Flageul et al., submitted-a). Moreover, a rapid selection of genetic variants present in an IBV-Ark type vaccine was already observed after one passage in chickens (McKinley et al., 2008; van Santen and Toro, 2008).

From a unique egg passage virus stock, the evolution of TCoV-inoculum was repeatable as evolution profiles were highly similar in the three replicates. Such results could indicate that natural evolution of an avian coronavirus is not random but could tend towards an optimal sequence adapted to the host. Also, this would mean that from a given viral population, no matter how many inoculations would be performed, the evolution of the virus would be the same as long as the replication conditions remain unchanged.

The impact of these mutations on protein functionality has been analysed and it seemed that the mutations determined in the polyprotein 1ab did not deeply change the polarity or the charge of the amino acid. Mutations in the spike protein were located in S1, the most variable part of the spike protein. Cch4 replaced a Serine (polar uncharged) by an Isoleucine (hydrophobic). The two amino acids have different biochemical properties and because this protein position is located at the surface of the protein, when compared to IBV spike protein (Shang et al., 2018), the modification could potentially modify the binding of neutralizing antibodies or the binding of the spike protein to its receptor. Cch5 replaced a Serine by an Asparagine, two polar uncharged amino acids so this modification may not impact the protein conformation. This position is not located at the surface of the spike protein, which indicates that both amino acids cannot be targeted by neutralising antibodies. Thus, based only on protein functionality, the reasons for these mutations to be selected in all the replicates remains unknown. Possibly, the modified RNA sequence may allow a better replication of the viral population in turkeys but this hypothesis seems unlikely since the concentration of viral RNA in cloacal swabs taken from infected birds at the end of each passage was similar.

In conclusion, this work showed that five positions in the viral genome have been modified at the consensus level as soon as the first passage and were fixed until passage 10, which supports the hypothesis that TCoV, like its close relative IBV has the potential to generate variants quickly.

This factor adds to difficulty of controlling TCoV infections because quickly generated variants help it survive different environmental and physical pressures. In light of the difficulties in controlling IBV despite numerous vaccines and vaccine programmes and the ongoing question of whether live attenuated vaccination is helping drive IBV evolution, it raises an important question of whether development of live attenuated TCoV vaccines are in the best interest of the turkey.

Table 1: Concentration of viral RNA in cloacal swabs at each passage

Replicates	Passages									
	1	2	3	4	5	6	7	8	9	10
A	4.95	5.37	2.08	4.63	4.41	4.36	4.18	3.52	4.87	4.15
B	4.51	4.45	4.46	3.82	3.70	3.70	4.11	5.08	3.87	4.84
C	4.93	4.73	4.58	4.60	4.76	4.69	5.09	4.79	4.63	3.84

log10 (viral RNA copy number/ μ L)

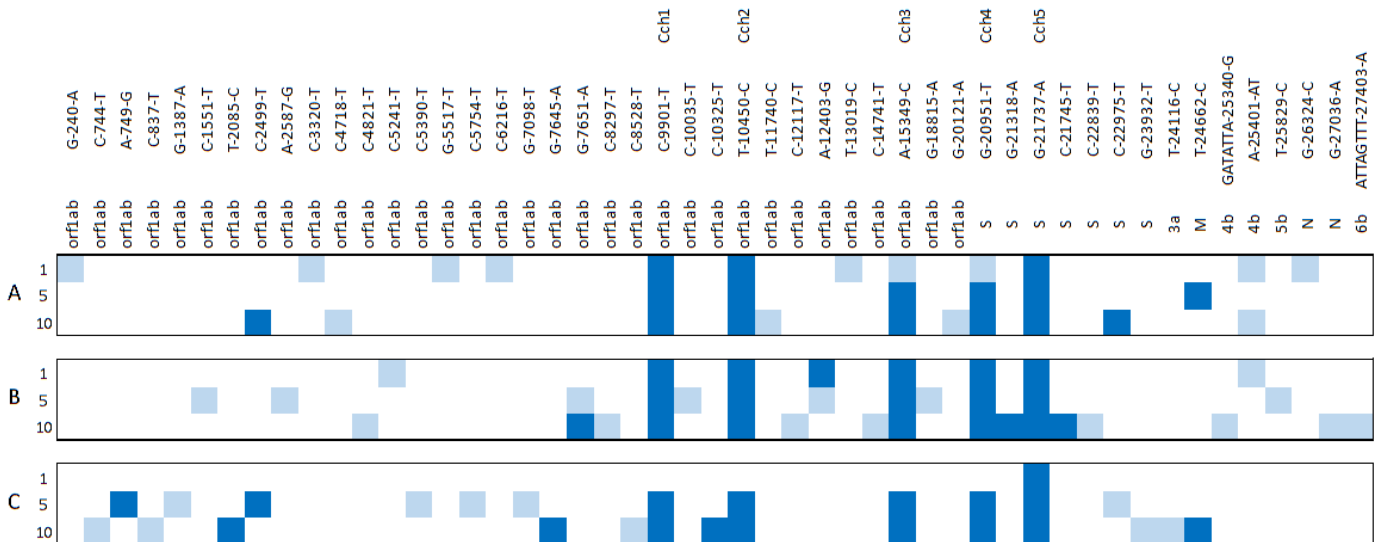


Figure 1: Evolution of Fr-TCoV during passages in turkeys. Evolution was assessed using NGS after passage 1, 5 and 10 and NGS data were processed through the Viral Variant Visualiser pipeline. Subjects are presented in lines A, B and C and detected genetic variants are presented in lines. Genomic regions, genome positions and predicted protein modifications are indicated above the graphic. If a genetic variant is present in a subject, a blue box is set: light blue colour for variants detected below the consensus level or dark blue colour for variants detected at the consensus level. Cch = Consensus change analysed in the text.

References

- Barnes, H.J., Guy, J.S., Vaillancourt, J.P., 2000. Poults enteritis complex. *Rev. - Off. Int. Epizoot.* 19, 565–588.
<https://doi.org/10.20506/rst.19.2.1234>
- Brown, P.A., Courtillon, C., Weerts, E.A.W.S., Andraud, M., Allée, C., Vendembeuche, A., Amelot, M., Rose, N., Verheije, M.H., Etteradossi, N., 2018. Transmission Kinetics and histopathology induced by European Turkey Coronavirus during experimental infection of specific pathogen free turkeys. *Transbound Emerg Dis* 66, 234–242. <https://doi.org/10.1111/tbed.13006>
- Callison, S.A., Hilt, D.A., Boynton, T.O., Sample, B.F., Robison, R., Swayne, D.E., Jackwood, M.W., 2006. Development and evaluation of a real-time Taqman RT-PCR assay for the detection of infectious bronchitis virus from infected chickens. *J. Virol. Methods* 138, 60–65. <https://doi.org/10.1016/j.jviromet.2006.07.018>
- Flageul, A., Allée, C., Courtillon, C., Béven, V., Quenault, H., Blanchard, Y., Amelot, M., Courtois, D., De Wit, S.J.J., Etteradossi, N., Grasland, B., Brown, P.A., submitted-a. Avian coronavirus: the genomic evolution of the infectious bronchitis virus, differs in chickens after vaccination.
- Flageul, A., Courtillon, C., Allée, C., Leroux, A., Blanchard, Y., Deleforterie, Y., Grasland, B., Brown, P.A., submitted-b. A modified in-ovo method for isolating Turkey Coronaviruses yields better Next Generation Sequencing results.
- Flageul, A., Lucas, P., Hirchaud, E., Touzain, F., Blanchard, Y., Etteradossi, N., Brown, P., Grasland, B., 2021. Viral variant visualizer (VVV): A novel bioinformatic tool for rapid and simple visualization of viral genetic diversity. *Virus Research* 291, 198201. <https://doi.org/10.1016/j.virusres.2020.198201>
- Gomaa, M.H., Yoo, D., Ojkic, D., Barta, J.R., 2008. Seroprevalence of turkey coronavirus in North American turkeys determined by a newly developed enzyme-linked immunosorbent assay based on recombinant antigen. *Clin. Vaccine Immunol.* 15, 1839–1844. <https://doi.org/10.1128/CVI.00319-08>
- Guionie, O., Courtillon, C., Allee, C., Maurel, S., Queguiner, M., Etteradossi, N., 2013. An experimental study of the survival of turkey coronavirus at room temperature and +4°C. *Avian Pathol.* 42, 248–252.
<https://doi.org/10.1080/03079457.2013.779364>
- McKinley, E.T., Hilt, D.A., Jackwood, M.W., 2008. Avian coronavirus infectious bronchitis attenuated live vaccines undergo selection of subpopulations and mutations following vaccination. *Vaccine* 26, 1274–1284.
<https://doi.org/10.1016/j.vaccine.2008.01.006>
- Shang, J., Zheng, Y., Yang, Y., Liu, C., Geng, Q., Luo, C., Zhang, W., Li, F., 2018. Cryo-EM structure of infectious bronchitis coronavirus spike protein reveals structural and functional evolution of coronavirus spike proteins. *PLoS Pathog.* 14, e1007009. <https://doi.org/10.1371/journal.ppat.1007009>
- van Santen, V.L., Toro, H., 2008. Rapid selection in chickens of subpopulations within ArkDPI-derived infectious bronchitis virus vaccines. *Avian Pathol.* 37, 293–306. <https://doi.org/10.1080/03079450802043783>

8.2 Evolution de l'IBV en absence ou présence d'une immunité vaccinale pré-existante

Dans cette étude, l'évolution de la souche IBV-D388, stock viral issu de quatre passages sur œufs de poule, a été caractérisée par NGS au cours de trois passages du virus sur des poulets ayant ou non été préalablement vaccinés avec un vaccin vivant atténué IBV-H120 de type Mass (essai hétérologue IBV H120-D388). En parallèle, afin de vérifier l'efficacité de ce vaccin, un second essai a été réalisé en utilisant la souche IBV-M41 comme virus épreuve (essai homologue IBV H120-M41), selon le même protocole que pour l'essai H120-D388 (décrit ci-dessous).

Les expérimentations sur animaux ont été réalisées dans les installations de biosécurité de niveau 2 du laboratoire ANSES de Ploufragan-Plouzané-Niort, conformément aux réglementations européenne et française sur le bien-être animal. Les poulets exempt d'organismes pathogènes spécifiés (EOPS) provenaient du service d'élevage et d'expérimentation avicole du laboratoire ANSES de Ploufragan-Plouzané-Niort. Le protocole expérimental décrit ci-dessous a été approuvé (Référence APAFIS # 5178-2016041118128688 v5) par le comité d'éthique Anses / ENVA / UPEC enregistré sous le numéro 16 et par le ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation.

Trente poulets White Leghorn EOPS d'un jour d'âge (1 j.a.) ont été séparés en deux groupes de quinze oiseaux et chaque groupe a été placé dans un isolateur d'élevage. Le premier groupe (oiseaux vaccinés (VB)) a été inoculé à 1 j.a. par voie oculaire avec un vaccin vivant atténué IBV-H120 préparé dans du milieu STAB. Le deuxième groupe (oiseaux non vaccinés (NVB)) a été inoculé de la même manière avec le milieu STAB. Pour le premier passage dans chaque groupe, cinq oiseaux de 24 j.a., ont été déplacés vers un nouvel isolateur et éprouvés par voie oculaire avec 10^4 DIE50 de virus épreuve IBV-D388 (nommé inoculum D388). À 29 j.a. les oiseaux éprouvés ont été euthanasiés par électronarcose suivie d'une saignée. Sur ces oiseaux, la trachée et les reins ont été prélevés et individuellement broyés avant d'être centrifugés. Les surnageants, qui représentent les stocks de virus issus de la trachée et des reins du premier passage de IBV-D388 dans le groupe VB et dans le groupe NVB, ont été individuellement récoltés et étiquetés avec le numéro d'oiseau suivi de D388_{pass1Tr}VB ou D388_{pass1Tr}NVB pour les prélèvements de trachées ou de D388_{pass1Kd}VB et D388_{pass1Kd}NVB pour les prélèvements de reins et stockés à -80°C. Avant surgélation, 200 µL de chacun des cinq stocks de D388_{pass1Tr}VB et 200 µL de chacun des cinq stocks de D388_{pass1Kd}VB ont été mélangés ensemble pour créer un stock de virus groupé D388_{pass1TrKd}VB. De même, 200 µL de chacun des cinq stocks D388_{pass1Tr}NVB et 200 µL de chacun des cinq stocks D388_{pass1Kd}NVB ont été mélangés ensemble pour créer un stock de virus groupé D388_{pass1TrKd}NVB. Deux autres passages ont été réalisés selon le même protocole, avec à

chaque fois un étiquetage des stocks de virus individuels et des stocks de virus groupés qui a suivi le même format. Les animaux du passage 2 ont été éprouvés avec D388_{pass1TrKd}VB ou D388_{pass1TrKd}NVB selon leur groupe d'appartenance, et les animaux du passage 3 ont été inoculés avec D388_{pass2TrKd}VB ou D388_{pass2TrKd}NVB également selon leur groupe d'appartenance. Ce protocole a également été suivi en utilisant le virus épreuve Mass41.

Au cours de l'expérimentation IBV homologue H120-M41, une réplication virale a été détectée lors des trois passages dans le groupe NVB. Aucune réplication virale n'a été détectée dans le groupe VB, à l'exception d'un sujet lors du premier passage. En conséquence, aucune comparaison évolutive de la souche M41 en fonction du statut immunitaire des sujets n'a pu être effectuée. Cependant, cette expérimentation permet de montrer que la vaccination H120 était fonctionnelle.

Au cours de l'expérimentation IBV hétérologue H120-D388, une réplication virale a été détectée lors de tous les passages, dans les reins de tous les sujets du groupe VB et du groupe NVB et dans la trachée d'une majorité de sujets des deux groupes. Dans le groupe NVB, les sujets du troisième passage présentaient plus de lésions rénales (néphrite) que les sujets des passages 1 et 2, alors qu'aucune néphrite n'a été observée dans le groupe VB.

Dans le but de procéder à des analyses de variants génétiques, les broyats de trachée et de reins ont été inoculés une fois sur œufs de poule embryonnés afin d'amplifier les populations virales présentes dans les échantillons expérimentaux et ainsi de les séquencer plus facilement en NGS. Après traitement des données NGS avec le pipeline nouvellement développé, des variants ont été détectés dans les groupes NVB et VB et une sélection de certains variants spécifiques de chaque groupe a été observée, provoquant un changement de la séquence consensus de l'inoculum D388 (Tableau 3).

Dans le groupe NVB, il y a plus de variants génétiques détectés (58 dans la trachée et 50 dans les reins) que dans le groupe VB (36 dans la trachée et 39 dans les reins). Par ailleurs, il y a plus de variants à avoir été sélectionnés dans le groupe NVB (Var1 à 7) que dans le groupe VB (Var8) (Tableau 3). Ces résultats peuvent être expliqués par des charges génomiques virales légèrement supérieures dans le groupe NVB que dans le groupe VB, traduisant un niveau de réplication virale plus faible dans le groupe VB réduisant ainsi les chances d'évolution ou d'identification de la population virale. Cependant, certains des variants sélectionnés dans le groupe NVB ont également été détectés dans le groupe VB mais sans être sélectionnés. Ceci pourrait donc indiquer que, dans ces conditions expérimentales, la vaccination a empêché la

sélection de ces mutations. Ces différents résultats peuvent donc indiquer que l'évolution de l'IBV pourrait dépendre du statut immunitaire des oiseaux vis-à-vis de l'IBV.

Tableau 3: Variants génétiques identifiés après sélection au cours des passages sur poulets lors de l'expérimentation IBV hétérologue H120-D388

Groupe	Variant	Position génomique	Consensus de l'inoculum	Nucléotide(s) sélectionné(s) après passage	Région protéique	Impact protéique
NVB	Var1	10673	C	T	PP1ab - NSP7	synonymous
NVB	Var2	10994	T	G	PP1ab - NSP8	H_3520_Q
NVB	Var3	15707	T	G	PP1ab - NSP13	F_5092_V
NVB	Var4	20328	TTGATTCTGATA ^{del}		Spike - S1	FDSDN_20_Y
NVB	Var5	20458	ATGT	GTGA	Spike - S1	HV_64_RE
NVB	Var6	23736	T	G	Spike - S2	*_1157_E
NVB	Var7	25040	C	T	Membrane	synonymous
VB	Var8	11624	C	T	PP1ab - NSP9	synonymous

Dans chaque groupe, les mêmes sélections de variants ont été observées dans la trachée et dans les reins. Ceci est probablement le résultat d'un biais expérimental car à la fin de chaque passage, les homogénats de trachée et de reins ont été mélangés ensemble afin de constituer l'inoculum du passage suivant. Mais ce résultat a montré que les virus se répliquant dans la trachée étaient également capables de se répliquer dans le rein et vice versa, mettant en évidence le manque d'informations concernant le tropisme tissulaire de l'IBV.

Dans les deux groupes, la sélection de variants a été rapide, car un seul passage a suffi pour que les variants sélectionnés modifient la séquence consensus. Une telle rapidité de sélection a déjà été observée avec un vaccin de type IBV-Ark. Dans ce vaccin, des variants génétiques ont été identifiés et certains ont été sélectionnés après un seul passage chez le poulet, modifiant la séquence consensus du génome vaccinal (McKinley et al., 2008; van Santen et Toro, 2008). Le même scénario est retrouvé ici puisque quatre des variants sélectionnés (Var2, Var3, Var4 et Var 6) étaient présents dans la population virale de l'inoculum D388. Par ailleurs, Var5 (ATGT-20458-GTGA) n'était pas présent dans l'inoculum D388. Cependant le dernier nucléotide de Var5 (correspondant à la mutation T-20461-A) était un variant présent dans la population virale de l'inoculum D388 et qui a ensuite été intégré à Var5 après un passage chez le poulet. Mais parmi les variants sélectionnés, certains n'ont pas été détectés dans la population virale de l'inoculum D388. Cela peut signifier que: i) les variants sont apparus pendant les passages sur poulet et ont ensuite été sélectionnés au niveau du consensus, ou bien que ii) les variants étaient présents dans la population virale de l'inoculum D388, mais en dessous du seuil de détection de 7%.

Parmi les variants sélectionnés, seul Var2, Var3, Var4, Var5 et Var6 ont provoqué des changements de séquences des protéines encodées.

Var2 et Var3 ont provoqué une modification de séquences des protéines NSP8 et NSP13 respectivement. En comparant avec les séquences protéiques présentes dans la base de données du NCBI, les séquences protéiques variantes correspondaient aux séquences les plus fréquemment retrouvées dans cette base de données. À l'inverse, les séquences protéiques de l'inoculum D388 n'étaient pas (Var2), ou très peu (Var3), présentes dans les séquences d'IBV de la base de données. La fonctionnalité des protéines NSP8 et NSP13 après trois passages sur poulets ne semble donc pas impactée.

Var4 et Var5 ont provoqué des modifications de séquence de la protéine de spicule, dans le domaine S1. Mais les changements de nucléotides du gène S n'ont pas permis de modifier la classification génotypique du virus variant basée sur S1 : l'inoculum D388 et le virus variant appartiennent tous les deux à la lignée GI-19.

Var6 entraîne la suppression du codon STOP de la protéine de spicule de l'inoculum D388, permettant l'ajout de neuf acides aminés (EQYRPKKS~~V~~-stop) à son extrémité C-terminale. Dans cette extension protéique, se trouve un motif di-lysine (KKxx) correspondant à un signal de rétention des protéines cellulaires dans le réticulum endoplasmique (Ma et Goldberg, 2013). Ce motif a déjà été identifié dans la protéine S de plusieurs d'IBV, dont la souche IBV-Beaudette. Par ailleurs, au cours d'une expérience *in-vitro*, ce motif issu de IBV-Beaudette a été ajouté à l'extrémité C-terminale de la glycoprotéine de surface du virus de la stomatite vésiculaire, qui est normalement adressée à la membrane plasmique après sa production. Cet ajout empêchait l'adressage normal de la glycoprotéine à la membrane plasmique, en la maintenant dans le compartiment intermédiaire entre le réticulum endoplasmique et l'appareil de Golgi (ERGIC) (Lontok et al., 2004). Or le bourgeonnement des particules virales d'IBV est effectué au niveau du ERGIC, et dans la population virale de l'inoculum D388, Var6 est un variant représentant 20% des séquences virales, ce qui implique que la majorité des protéines de spicules ne contiennent pas le motif di-lysine. Par conséquent, comment la majorité des protéines de spicule de l'inoculum D388 serait retenue dans l'ERGIC? Dans les bases de données, les deux types de séquences du gène S, codant ou non l'extension protéique, sont retrouvées. D'autres motifs d'acides aminés pourraient également être impliqués dans la rétention de la protéine de spicule dans le ERGIC, tel qu'un motif tyrosine (Winter et al., 2008).

Pour conclure, ce travail a révélé que l'évolution de l'IBV était très rapide et qu'elle était différente selon que les sujets avaient été vaccinés ou non.

Ce travail a fait l'objet d'un article qui sera soumis à *Journal of Virology*.

Avian coronavirus: the genomic evolution of the infectious bronchitis virus differs in chickens after vaccination.

Alexandre Flageul¹, Chantal Allée¹, Céline Courtillon¹, Véronique Béven², Hélène Quenault², Yannick Blanchard², Michel Amelot³, David Courtois³, Sjaak De Wit⁴, Nicolas Eterradossi⁵, Béatrice Grasland¹ and Paul Alun Brown¹.

VIPAC¹, GVB², SELEAC³ unit, French Agency for Food, Environmental and Occupational Health & Safety (ANSES), Laboratoire de Ploufragan-Plouzané-Niort, Université de Rennes 1

⁴GD Animal Health, Deventer, The Netherlands

⁵MD French Agency for Food, Environmental and Occupational Health & Safety (ANSES), Université de Rennes 1

Abstract

Infectious Bronchitis virus (IBV) is a long time known avian coronavirus that causes significant economic losses for the chicken industry. Despite the fact that several IBV vaccines are available and widely used around the world, novel IBV strains continue to be characterized. Some authors have previously suggested that vaccine inducing protection, or partial protection, against IBV, may contribute to the emergence of new IBV strains. In order to study this hypothesis, three successive passages of a challenge IBV were made three weeks after having vaccinated or sham inoculated chickens at day of age. The evolution of full-length viral genomes replicating in the tracheas and kidneys of these birds was determined over passages by high-throughput sequencing and processing data with a newly developed bioinformatics tool that was designed to detect viral genetic variations that compose the minority of the viral population. All sequenced samples were compared to the inoculum sequence.

In the unvaccinated group, a rapid but progressive selection of seven genetic variants was observed over the three passages, finally modifying the majority of the viral population and the consensus sequence. Among them, five were non-synonymous modifying the protein sequences. In the vaccinated group, one genetic variant was selected over the three passages and this modification was absent from the unvaccinated group. This nucleotide modification was synonymous. Interestingly, within each group, the same selection occurred in tracheal and kidney tissues. This work revealed that under these experimental conditions an IBV genome population evolved rapidly and completely differently in birds with different IBV immune status.

Introduction

The Coronaviridae is a viral family characterised by a large single-stranded positive-sense RNA genome. In this family, two sub-families are described the Letovirinae and the Orthocoronavirinae. In the latter, four genus have been established, including the Gammacoronavirus genus in which the Infectious Bronchitis Virus (IBV) is the prototype virus.

IBV was first isolated in the USA in 1933 by Bushnell and Brandly, from chickens with Infectious Bronchitis (IB), a respiratory disease of chickens initially described in 1931 by Schalk and Hawn (Delaha et al., 1954). Since the initial isolation, many different genotypes and antigenic types have been discovered around the world (Jackwood, 2012; Valastro et al., 2016) and some of these produce clinical manifestations other than or in addition to respiratory, including nephritis and false layer syndrome. All continue to cause serious economic losses for the poultry industry (Colvero et al., 2015). To date, no anti-viral treatments exists for IBV and so infection is controlled through vaccination with live-attenuated and inactivated viruses, and by strict attention to hygiene, such as complete cleaning and disinfection of animal houses.

To date the live-attenuated vaccines are the primary source of protection and the inactivated vaccines serve as a support to boost the immune response. However, due to the diversity of IBV strains around the world, no single live-attenuated vaccine confers protection against all strains. This has led to the development of an impressive amount of live-attenuated vaccines derived from different IBV strains and their use in many different vaccination programs. For instance, in Poland over a period of twenty months, more than 70 vaccination protocols were inventoried, including the simultaneous administration of six live-attenuated vaccines (Legnardi et al., 2019). This heavy use of live-attenuated vaccines has been suggested as a potential factor contributing to the emergence of new IBV strains either through the introduction of IBV vaccine strains that were not previously indigenous to the geographic region, through increased chances of recombination or pressure / selection imposed by pre-existing immunity (Jackwood et al., 2012; Toro et al., 2012). Concerning the effects of pre-existing immunity on IBV, full genome evolution through pressure / selection there is currently no experimental data available.

Thus, this question is addressed in the current study. Chickens with or without a pre-existing immunity to a Massachusetts type live-attenuated vaccine (strain H120) were used to perform three passages of a vaccine-homologous Mass41 IBV isolate or three passages of a vaccine-heterologous D388 IBV isolate. The full-length genomes of passaged viruses were sequenced using Next Generation Sequencing (NGS) to compare their evolution in chickens with different IBV immune status.

Material and methods

Viruses and vaccine

Lyophilized D388 IBV was re-suspended in 1 ml of stabilizing buffer (STAB, composed with Saccharose [20 g/L], KH₂PO₄ [0.52 g/L], NaHPO₄·12H₂O [2.58 g/L]). Re-suspended D388 and a virus stock of the strain M41 were titrated in eggs (Kint et al., 2015), using the Reed and Muench method (Reed and Muench, 1938). D388 was diluted in STAB to adjust the viral titer to 10⁵ egg infectious dose (EID) 50 / mL. M41 titer was 10^{5.33} EID₅₀ / mL and was not diluted. Lyophilized H120 vaccine was re-suspended in 100 mL of STAB, according to the manufacturer recommendation so that one dose of vaccine equals 50µL.

Preparation of embryonic Chicken Kidney Cells (eCK)

Chicken embryos of 19 days of age (d.o.a.) were euthanized. Then, kidneys were collected and washed with an excess of Phosphate Buffer Saline (PBS) at 37°C for three minutes in an Erlenmeyer flask with gentle agitation. After centrifugation at 375g for five minutes (min) at room temperature, the supernatant was carefully removed and excess of PBS supplemented with trypsin (0.05 %), ethylenediaminetetraacetic acid (EDTA, 0.1%), penicillin (100 units / ml) and streptomycin (100 µg /ml) was added to dissociate the cells. Then the suspension was incubated for one min at 37°C again with gentle agitation. After centrifugation at 375g for five min at room temperature the supernatant was carefully removed and the dissociating step was repeated a second time. Then, foetal calf serum was added (10% final) and centrifuged at 375g for five min at room temperature. Supernatant was removed and cells were resuspended in a minimum of BHK medium (tryptose phosphate broth, Glasgow Modified Eagle Medium, penicillin (100 units / ml) and streptomycin (100 µg /ml)). Cell suspension was then filtered over six layers of sterile gauze band and cell concentration was adjusted to 5.10⁵ cells per millilitre.

Experimental trial

All experiments were performed in accordance with EU and French regulations on animal welfare in experiments. The protocol was approved (Reference APAFIS#5178-2016041118128688 v5) by the ANSES/ENVA/UPEC ethical committee registered under number #16 and by the French ministry of Higher education, research and innovation.

Two experimental trials were carried out, i) **H120-D388** (vaccination and challenge with vaccine heterologous D388) and ii) **H120-M41** (vaccination and challenge with vaccine homologous M41).

i) Trial H120-D388

Thirty specific pathogen free (SPF) one day-old White Leghorn chickens were separated into two groups of fifteen birds and each group was placed in a rearing isolator in ANSES' biosafety level 2 facilities. The first group (vaccinated birds (VB)) was vaccinated at 1 d.o.a. by eye drop with 60 μ L of H120 live-attenuated vaccine. The second group (non-vaccinated birds (NVB)) was sham-inoculated in the same way with STAB medium. At 24 d.o.a., blood samples were taken from all birds to check their IBV immune status. For VB and NVB, five animals at 24 d.o.a., were moved to a new isolator and then challenged with 10⁴ egg infectious dose 50 (10⁴ EID₅₀) by eye drop with D388. At 29 d.o.a., challenged animals were euthanized by electronarcosis followed by bleeding. Blood, tracheal and kidney tissues were then collected from these birds. Tracheal and kidney tissues were individually homogenized in STAB medium (1 gram/1mL of STAB) and centrifuged (1000g, 10 min, 4°C). Pellets were discarded and supernatants collected. These supernatants were stored individually at -80°C and represented tracheal and kidney virus stocks of first passage D388 in VB and in NVB. Stocks were labelled with the bird number followed by D388_{pass1Tr}VB or D388_{pass1Tr}NVB for the trachea and with the bird number followed by D388_{pass1Kd}VB and D388_{pass1Kd}NVB for the kidneys. Two hundred microliters from each of the five D388_{pass1Tr}VB stocks and 200 μ L from each of the five D388_{pass1Kd}VB stocks were mixed together to create a pooled virus stock D388_{pass1TrKd}VB. Equally, 200 μ L from each of the five D388_{pass1Tr}NVB stocks and 200 μ L from each of the five D388_{pass1Kd}NVB stocks were mixed together to create a pooled virus stock D388_{pass1TrKd}NVB.

At 29 d.o.a a new group of five VB and five NVB were moved from the rearing isolators to new isolators and challenged with D388_{pass1TrKd}VB or D388_{pass1TrKd}NVB respectively. After challenge, the process described above for passage one was repeated. This yielded individual virus stocks that were labelled with the bird number followed by, D388_{pass2Tr}VB, D388_{pass2Tr}NVB, D388_{pass2Kd}VB or D388_{pass2Kd}NVB and two pooled virus stocks: D388_{pass2TrKd}VB and D388_{pass2TrKd}NVB.

At 36 d.o.a the final groups of five VB and five NVB were challenged with pooled virus stocks D388_{pass2TrKd}VB or D388_{pass2TrKd}NVB respectively in the rearing isolators. After challenge, the process described above for passage one and two was repeated. This yielded individual virus stocks that were labelled with the bird number followed by: D388_{pass3Tr}VB, D388_{pass3Tr}NVB, D388_{pass3Kd}VB or D388_{pass3Kd}NVB.

Blood samples were taken at each passage just before challenge to check IBV immune status and all viral stocks were tested for the presence of viral RNA using an avian coronavirus RT-qPCR as previously described (Callison et al., 2006). Clinical signs (sneezing, rales, prostrations) were

monitored for the duration of the trial and during necropsies (for collection of tracheas and kidneys) macroscopic lesions were noted.

ii) Trial H120-M41

The only change from the trial described above was the use of M41 as challenge virus. Virus samples were coded in the same way, the only difference was the replacement of D388 by M41. As an example: the stock D388_{pass1Tr}VB of trial H120-D388 was labelled M41_{pass1Tr}VB in this trial.

IBV Enzyme-Linked-Immuno-Sorbent-Assays (ELISA)

Blood samples were centrifuged (1160 g, 15 min, 4°C) and sera were tested with commercial kit IBV CK119 (Biochek) according to the manufacturer's instructions. Results were interpreted as recommended, with a positivity threshold of absorbance of 0.2.

Preparation of viral samples for NGS

All samples plus D388 inoculum were passaged one time in embryonated SPF chicken eggs to optimize samples for NGS. This system of optimization was chosen as it has been promoted as an efficient method in producing high levels of NGS data from IBV isolates (Keep et al., 2015).

For this procedure, individual stocks were diluted (1/10) in PBS supplemented with penicillin (10 U/μL), streptomycin (10 μg/μL) and amphotericin B (250 μg/μL) and incubated for one hour at 4°C. These were then inoculated into the allantoic cavity of nine day old embryonated SPF chicken eggs. Eggs were incubated for 6 days at 37°C and candled every day to check for embryo mortality. Chicken embryos that died during the first 24 hours of incubation were discarded. At the end of the incubation period, or for embryos that died later than 24 hours post inoculation, allantoic liquids were harvested as previously described (Guy, 2015) then centrifuged at 1160 g for 15 min at 4°C to produce clarified inoculum and clarified virus samples.

RNA extraction

RNAs were extracted from the inoculum and all clarified virus samples using the King Fisher automatic extractor using magnetic beads, washing and elution buffers from NucleoMag Vet kit (Macherey-Nagel). Briefly, 200 μL of samples were lysed with 300 μL of RLT buffer supplemented with 3 μL of β-mercaptoethanol and 16 μg of glycogen. After the lysis step, 20 μL of magnetic beads and 400 μL of isopropanol were added for RNA precipitation. Three washing steps were realised: two using washing buffers and one using 80% ethanol. Finally, RNA was eluted using 100 μL of elution buffer.

Illumina sequencing and NGS data processing

RNA libraries were prepared using the NEBNext Ultra II Directional RNA library. Briefly, ribosomal RNA were removed with RNaseH and remaining RNA was converted into double strand DNA (dsDNA) by random reverse transcription and second strand DNA synthesis. Then, sequencing adaptors were ligated to dsDNA and a short PCR enrichment was performed before sequencing. NGS was performed using NextSeq Illumina technology. NGS data were processed twice with the Viral Variant Visualiser (VVV) pipeline (Flageul et al., 2021) to assess IBV variants. The first analysis involved all the reads in the NGS file while the second analysis removed reads considered as duplicates (identical reads). For reasons previously defined (Flageul et al., 2021), only genetic variants detected in both analysis and present in the viral population above a threshold value of 7% were analysed. Data were considered robust when the median coverage depth was greater than 500 and when a minimum of 200 reads was present in locations where variants were detected. Other data were annotated as “low coverage”.

***In silico* and *in vitro* analysis of variants with modified protein sequences**

PSI-blast analysis (Altschul et al., 1997; Schäffer et al., 2001), using defaults parameters and restricted to the avian coronaviruses, was performed to research homologous proteins in the database. Variants with modified S proteins were analysed by phylogenetics and tested in virus neutralization assay as described below.

Viral neutralization and immunofluorescence assays

The D388 IBV inoculum, one clarified D388_{pass3Tr}NVB, one D388_{pass3Kd}NVB, one clarified D388_{pass3Tr}VB and one D388_{pass3Kd}VB were used for viral neutralization tests. Viruses were serially diluted from 10⁻¹ to 10⁻⁵ and distributed into two duplicate 96 well plates (plate 1 and 2) using 75 µL of dilution / well. Then 75 µL of BHK medium was added to each well of plate 1 and this was incubated at room temperature for 1h. 75µL of anti-D388 polyclonal serum (diluted 1/27 in BHK medium) was added to each well of plate 2 which was then incubated at room temperature for 1h. Following the 1h incubation period, 125µL of eCK cells were added to the wells of the two plates. Both plates were then incubated for two days at 37°C with 1% CO₂. After the two-day incubation period, supernatants were removed and cells were washed three times with PBS. Cells were fixed for 20 min at -20°C with a mix of ice-cold ethanol/acetone (ratio 1:1). Ethanol/acetone was removed and cells were washed three times with PBS and were dried before storage at -20°C. For immunofluorescence assay, plates were frozen, washed three times with PBS and incubated for one hour at room temperature (around 21°C) with a mouse monoclonal antibody targeting the

IBV Membrane protein (mAb-IBV-25.1, Wageningen Bioveterinary Research, diluted 1/200 in PBS). After three washes with PBS, cells were incubated with the AlexFluor 488 goat anti-mouse IgG (H+L) (diluted 1/800 in PBS) for 45 min at room temperature. Cells were washed and stored in PBS at 4°C. Immunofluorescence was observed using FITC filter on an Olympus IX71 microscope with CoolLED pE-300 as source of light.

Phylogenetic analysis

The phylogenetic analysis of the 15-D388_{pass3Tr}NVB clarified virus sample was performed as described previously (Valastro et al., 2016). Briefly, a phylogenetic tree based on the S1 gene was realized using the Maximum Likelihood method based on the General Time Reversible model (Nei and Kumar, 2000) with a discrete Gamma distribution to model evolutionary rate differences among sites (4 categories (+G, parameter = 1,0432)). The rate variation model allowed for some sites to be evolutionarily invariable ([+I], 10.76% sites). Bootstrap sampling method was used over 1000 replicates. The resulted tree was drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 199 nucleotide sequences of S1 from the Valastro classification (Valastro et al., 2016) and two additional S1 nucleotide sequences from the current H120-D388 trial (D388 inoculum and 15-D388_{pass3Tr}NVB). A total of 1826 positions were present in the final dataset. Evolutionary analyses were conducted using MEGA7 software (Kumar et al., 2016).

Statistical analysis

Statistic tests (comparison of mean of two independent samples) were performed using Wilcoxon test.

Results

H120-Mass41 trial

Prior to vaccine/placebo inoculation, all birds were seronegative for IBV (data not shown).

In the NVB group, birds were all seronegative prior to the different challenges (Table 1). Throughout the trial, NVB challenged with M41 remained healthy and did not display any notable clinical signs. No gross lesions were observed during necropsies of any birds. Five days after challenge, viral RNA was detected in the trachea of all birds of all passages, while viral RNA was detected in the kidneys of two out of five, two out of five and four out of five birds at passage 1, 2 and 3 respectively.

In the VB group, two birds died suddenly before challenge and therefore they were not taken into account in the experimental results. Two out of four, four out of five and three out of four were seropositive on the day of challenge at passage 1, 2 and 3 respectively. Throughout the trial, VB challenged with M41 remained healthy and did not display any notable clinical signs. No gross lesions were observed during necropsies of any birds. Viral RNA was only detected in one bird and uniquely in its tracheal sample (at a Ct value of 23.07).

H120-D388 Trial

Prior to vaccine/placebo inoculation, all birds were seronegative for IBV (data not shown).

In the NVB group, birds were all seronegative prior to the different challenges (Table 2). Five days after the initial D388 challenge, one bird died suddenly. At necropsy, this bird had hypertrophied kidneys. During passage 2, another bird showed signs of prostration in the evening of the fourth day after challenge. The following morning this bird had died and at necropsy, had haemorrhagic liver, hypertrophied kidneys and petechial trachea. Clinical signs were not observed for any other animal throughout the duration of the trial. However, at necropsy all birds in passage 3 had nephritis whilst birds of passage 1 and 2 did not. Viral RNA was detected in tracheas and kidneys of all birds of all passages with the exception of one bird (passage 3) whose kidney sample was positive and tracheal sample was negative.

In the VB group, two out of five, two out of five and five out of five birds were seropositive on the day of challenge at passage 1, 2 and 3 respectively. Neither clinical signs nor gross lesions were observed on animals throughout the duration of the trial except for one bird at passage 2 that had nephritis and petechial trachea. This bird was seronegative before challenge. Viral RNA was detected in kidneys of all birds at all passages, while viral RNA was detected in tracheas of three out of five, five out of five and three out of five birds for passage 1, 2 and 3 respectively. In the

VB group, viral genomic loads in tracheas and kidneys was lower than those found in the same tissues of the NBV animals (Figure 1).

Comparison of D388 genome evolution in VB and NVB

This section only concerns D388 as only one VB bird of the H120-M41 trial had a tracheal sample that was found positive for M41 viral RNA, making comparative analysis on the evolution of this genome in VB and NVB impossible.

Sequencing of clarified D388 inoculum

The consensus sequence of clarified D388 inoculum was established from NGS data that generated a median coverage depth of 78396 (Figure 2) and a coverage depth of 497 when identical reads were removed. Both sets of data resulted in the detection of nine viral genetic variants with percentage of presence comprised between 7 and 25% (Figure 2). Among them, two were in the orf1ab at positions 10994 and 15707 (coding sequences of NSP8 and NSP13 respectively) and the seven others were detected in the S coding sequence, at positions 20328, 20461, 21654, 22126, 22417, 22921 and 23736.

Data processing and compilation

NGS data obtained from RNA extractions of the different clarified virus stocks and the clarified D388 inoculum were processed through the VVV pipeline (Flageul et al., 2021). VVV was programmed to compare D388 data from clarified virus stocks to the consensus sequence of the D388 inoculum. All detected variants were compiled into a single graphic which is shown in Figure 3. Each variant is indicated by a blue square which is light blue colour if the variant frequency was less than 50% of the total sequences at that position or dark blue colour if the variant was above the threshold of 50% of the total sequences, in other words meaning a consensus level change.

For all data, analysis involving all reads resulted in median coverage depths comprised between 5136 and 19647, except for one data obtained from an RNA extraction of D388_{pass1Tr}VB and another extraction of D388_{pass3Tr}VB which were 105 and 124 respectively.

Evolution of D388 genome population in NVB

A total of 58 variants were detected in the trachea and a total of 50 in the kidney. Seven of these variants (Var1 to 7) were selected over passage in both tissues and by the third passage had arrived at the consensus level in all birds except one which uniquely displayed the change in the trachea.

Var1 was not detected in the genome population of the clarified D388 inoculum. Var1 first appeared at passage 1 in the kidneys of one bird and at the consensus level. At passage 2, this variant was detected in the trachea of four birds (three at the consensus level, one below). The same variant was detected in the kidneys in four birds at the consensus level. At passage 3, this variant was detected in the trachea of all five birds and in the kidneys in four birds at the consensus level.

Var2 was present at 9 % in the genome population of the clarified D388 inoculum. At passage 1, Var2 was detected in the trachea of three birds (one at the consensus level, two below) and in the kidneys of two birds (one at the consensus level, one below). At passage 2, this variant was detected in the trachea of four birds (three at the consensus level, one below). The same variant was detected in the kidneys in four birds at the consensus level. At passage 3, this variant was detected in the trachea and kidneys of all five birds at the consensus level.

Var3 was present at 8 % in the genome population of the clarified D388 inoculum. At passage 1, Var3 was detected in the trachea of two birds (one at the consensus level, one below) and in the kidneys of one bird at the consensus level. At passage 2, this variant was detected in the trachea of four birds (three at the consensus level, one below). The same variant was detected in the kidney in four birds at the consensus level. At passage 3, this variant was detected in the trachea and kidneys of all five birds at the consensus level.

Var4 was present at 10 % in the genome population of the clarified D388 inoculum. At passage 1, Var4 was detected in the kidney of one bird at the consensus level. At passage 2, this variant was detected in the trachea of all five birds (four at the consensus level, one below). The same variant was detected in the kidney in four birds at the consensus level. At passage 3, this variant was detected in the trachea and kidneys of all five birds at the consensus level.

Var5 was not detected in the genome population of the clarified D388 inoculum. Var5 first appeared at passage 1 in the kidneys of one bird and at the consensus level. At passage 2, this variant was detected in the trachea of five birds (four at the consensus level, one below). The same variant was detected in the kidneys in four birds at the consensus level. At passage 3, this variant was detected in the trachea and the kidneys of all five birds at the consensus level.

Var6 was present at 20 % in the genome population of the clarified D388 inoculum. At passage 1, Var6 was detected in the trachea of three birds (below the consensus level) and in the kidneys of

three birds (two at the consensus level, one below). At passage 2 and 3, this variant was detected in the trachea and kidneys of all five birds at the consensus level.

Var7 was not detected in the genome population of the clarified D388 inoculum. Var7 first appeared at passage 1 in the kidneys of one bird and at the consensus level. At passage 2, this variant was detected in the trachea of five birds (four at the consensus level, one below). The same variant was detected in the kidneys in four birds at the consensus level. At passage 3, this variant was detected in the trachea and the kidneys of all five birds at the consensus level.

Evolution of D388 genome population in VB

A total of 36 variants were detected in the trachea and a total of 39 in the kidney. One of these variants (Var8) was selected over passage in both tissues.

Var8 was not detected in the genome population of the clarified D388 inoculum. Var8 first appeared at passage 1 in the kidneys of one bird and at the consensus level. At passage 2, this variant was detected in the trachea of all five birds at the consensus level. The same variant was detected in the kidneys in four birds (three at the consensus level, one below). At passage 3, this variant was detected in the trachea of two out of three birds at the consensus level and in the kidneys of four out of five birds at the consensus level.

Predicted protein impact of the non-synonymous selected mutations (NVB)

Of the eight variants described above, Var2, 3, 4, 5 and 6 were predicted to modify the amino acid (aa) sequence. Some of these were a result of an impact on a particular codon others were the result of deletion modifications.

Var2 (nucleotide (nt) mutation T-10994-G) modified the aa sequence (His-3520-Gln) of PP1ab in the NSP8 region. Var2 had replaced one polar positively charged aa by one polar uncharged aa. After two iterations of PSI-blast, 556 sequences of avian Gammacoronavirus NSP8 were collected and within these sequences the Glutamine of the variant was found in 555 sequences and the Histidine was not present (data not shown).

Var3 (nt mutation T-15707-G) modified the aa sequence (Phe-5092-Val) of PP1ab in the NSP13 region. Var3 had replaced one bulky hydrophobic aa by another of smaller size. After two iterations of PSI-blast, 438 sequences of avian Gammacoronavirus NSP13 were collected and within these sequences the Valine of the variant viruses was found in 413 sequences and Phenylalanine was present in four sequences (data not shown).

Investigation of Var4 and Var5 were conducted together as they both took place in the S1 region of the spike protein (upstream and in the middle of the hypervariable region 1 respectively). Var4

had the deletion TTGATTCTGATA (20328-20340) that removed aa 20 to 24 (FDSDN) and resulted in the creation of aa Y at position 20. Var5 (nt mutation ATGT-20458-GTGA) modified the aa sequence (HV-64-RE), replacing one positively charged aa and one hydrophobic aa by two positively charged aa. Because of these substantial modifications in the spike protein it was hypothesised that they may result in a classification of the variant viruses apart from the D388 inoculum and produce viruses that could not be neutralized by antibodies raised against the D388 inoculum. Classification was assessed using the previously described S1 classification method of IBV (Valastro et al., 2016) and included the S1 sequence of D388 inoculum clarified virus sample, and of the 15-D388_{pass3Tr}NVB clarified virus sample. Only one S1 sequence of the ten passage 3 viruses was analysed as all variants possessed Var4 and Var5. It was observed that both inoculum and the representative variant virus belonged to the GI-19 lineage (Figure 4). Neutralization studies showed that D388 was able to infect eCK cells in the absence of antibodies raised against the D388 inoculum yet was not able to infect eCK cells after prior adsorption with the anti-D388 polyclonal serum. The same study showed that variants D388_{pass3Tr}NVB and D388_{pass3Kd}NVB were not able to infect eCK cells under any condition. To control whether this impossibility of infection could be attributed to Var4 and Var5, the same neutralization assay was performed using D388_{pass3Tr}VB and D388_{pass3Kd}VB in which Var4 and Var5 were absent. Again, these variants viruses were not able to infect eCK cells under any condition.

Var6 (nt mutation T-23736 -G) replaced the stop codon of the spike protein by a glutamic acid and resulted in an extension of the spike protein by nine aa (stop-1157-EQYRPKKS_Vstop). This extended version of the spike protein has previously been deposited in GenBank databases for most IBV (444 out of 483 analysed S gene of IBV sequences).

Discussion

Since the initial isolation of IBV in the USA during the 1930's, and despite global vaccination of chicken flocks, numerous IBV strains are still detected around the world. Previous studies have suggested that vaccination could be one of the driving factors of IBV genome evolution (Jackwood et al., 2012; Toro et al., 2012). In the current study, two trials were performed to compare the evolution of the genome population of IBV-D388 and IBV-M41 isolates over three passages in unvaccinated and vaccinated birds. In the H120-M41 trial, viral RNA was detected in all NVB, but was detected in only VB at passage 1 demonstrating the efficacy of the homologous vaccination. Unfortunately this meant that no further analysis could be performed to compare the evolution of this virus in the different groups of birds.

Concerning H120-D388 trial, NVB subjects of passage 3 had nephritis, but subjects of passage 1 and 2 and subjects in VB did not, suggesting an increase of pathogenicity of D388 over passage in unvaccinated chickens. After processing NGS data with VVV (Flageul et al., 2021), a large number of variants were detected in both VB and NVB. Interestingly, only one variant was selected in VB upon passage and this variant was unique to this group, yet seven were selected upon passage in NVB. Of these eight variants, four were present in the clarified D388 virus inoculum and thus by default both groups of birds had the potential to select the same variants however, this did not occur. This may indicate that the seven variants selected by NVB were largely eliminated by the immunity induced by vaccination in VB. However it must be accepted that the slightly lower viral genomic loads in VB, could have reduced viral population diversity and this in conjunction with pre-existing immunity could have had an effect.

A second striking observation was the fact that within a group the same selected variants were observed in both tracheal and kidney tissues. This may have been a result of the nature of the inoculum being a mixture of tracheal and kidney homogenates from the previous passage however it is clear that these selected variants did not alter the capability of D388 to replicate in both tracheal and kidney tissues.

Three of the eight selected variants were not detected in the D388 inoculum viral population meaning that: i) they appeared during passages and then were selected at the consensus level, or ii) they were present in the viral population but below the 7% detection threshold. In addition, other variants present in the viral population of D388 inoculum were also detected in some birds at passage 1 before disappearing and many variants appeared sporadically in a limited number of birds before disappearing. These observations translate the capacity of the viral population to generate mutations that are continuously tested in a given environment and in which less fit mutations disappear.

In both groups, the selection of variants was fast, appearing at the first passage and at the consensus level. Such quick selection has also been observed in IBV-Ark type vaccine, in which existing genetic variants were selected and also reached the consensus level after only one passage in chickens (McKinley et al., 2008; van Santen and Toro, 2008).

Regarding the impact of new amino acid sequences coded by Var2 and Var3 on the functionality of NSP8 and NSP13 respectively, it was not assumed to have any affect as these amino acids were found at high percentage among Avian coronavirus sequences in the NCBI database. It would seem therefore that the variant changes represent a virus more adapted to the chicken than the parental D388 inoculum which had had previous successive replication in eggs to maintain viral stocks. However if this is the case then it is surprising that it was not selected in VB.

Var4 induced a protein modification (FDSDN-20-Y) that was not retrieved in the deduced amino acid sequences of Valastro et al. (2016) dataset. Replacement of the F by the Y amino acid has already been reported in 111 out of 199 Valastro sequences that did not belong to GI-19, but it was associated to other amino acids that partially or totally replaced the DSDN sequence.

Var6 modification led to an increased size of the spike protein by nine amino acids and added a di-lysine motif (KKxx) at the C-terminal extremity of the protein. Di-lysine motifs are known to be Endoplasmic Reticulum retention signals (Ma and Goldberg, 2013) and the addition of this motif from IBV-Beaudette strain to the Vesicular Stomatitis Virus Glycoprotein has been shown to maintain the recombinant protein in the intermediate compartment between the endoplasmic reticulum and the Golgi (ERGIC) (Lontok et al., 2004), which is the budding site of IBV. Most of the published IBV S-gene were predicted to code for the extended form of the spike protein and thus it would appear to be a preferred characteristic for the virus in the chicken. Interestingly though, although this variant was present in the inoculum, it was not selected or allowed to grow in VB.

To conclude, this work revealed that D388 inoculum displayed unusual amino acid sequences in place of Var2 and Var6, which might be the result of egg amplification. Moreover, this work revealed that under these experimental conditions an IBV genome population evolved rapidly and completely differently in vaccinated and non-vaccinated chickens. More variants were generated in non-vaccinated birds which most likely reflects the reduction of the viral population diversity through reduction of the viral load by the vaccine induced immunity. It is not known at this point what impact the different variants have on this strain or if they would be maintained upon further passage however it is clear that populations of viruses did not evolve the same in birds with different IBV immune status.

Acknowledgment

Authors thanks François-Xavier Briand for assistance with phylogeny analysis, the Institut Cerveau Moelle for having sequenced all the extracted RNAs, Houghton Trust and the French Group of World Veterinary Poultry Association for financing the participation to the WVPA congress. This work was partly funded by Agence National de Sécurité Sanitaire, de l'Environnement et du travail (ANSES), the "Région Bretagne", the "Université Bretagne Loire" and the "Université de Rennes" (UNIR) (code of convention : 180019).

Table 1: Numbers of seropositive birds and viral RNA qRT-PCR positive organs in the H120-M41 trial

	M41 NVB			M41 VB		
	Passage 1	Passage 2	Passage 3	Passage 1	Passage 2	Passage 3
seropositivity (1 doa)	0/5	0/5	0/5	0/5	0/5	0/5
seropositivity before Passage 1	0/5	0/5	0/5	2/4	3/5	1/4
seropositivity before Passage 2	/	0/5	/	/	4/5	/
seropositivity before Passage 3	/	/	0/5	/	/	3/4
qRT-PCR positive trachea	5/5	5/5	5/5	1/4	0/5	0/4
qRT-PCR positive kidney	2/5	2/5	4/5	0/4	0/5	0/4

Table 2: Numbers of seropositive birds and viral RNA qRT-PCR positive organs in the H120-D388 trial

	D388 NVB			D388 VB		
	Passage 1	Passage 2	Passage 3	Passage 1	Passage 2	Passage 3
seropositivity (1 doa)	0/5	0/5	0/5	0/5	0/5	0/5
seropositivity before Passage 1	0/5	0/5	0/5	2/5	2/5	3/5
seropositivity before Passage 2	/	0/5	/	/	2/5	/
seropositivity before Passage 3	/	/	0/5	/	/	5/5
qRT-PCR positive trachea	5/5	5/5	4/5	3/5	5/5	3/5
qRT-PCR positive kidney	5/5	5/5	5/5	5/5	5/5	5/5

Table 3: Summary of selected genetic variations over passage in the H120-D388 trial

	Genomic Position	Ref	Alt	Protein region	Protein impact
Var1	10673	C	T	PP1ab - NSP7	synonymous
Var2	10994	T	G	PP1ab - NSP8	H_3520_Q
Var3	15707	T	G	PP1ab - NSP13	F_5092_V
Var4	20328	TTGATTCTGATA ^{del}		Spike - S1	FSDSN_20_Y
Var5	20458	ATGT	GTGA	Spike - S1	HV_64_RE
Var6	23736	T	G	Spike - S2	*_1157_E
Var7	25040	C	T	Membrane	synonymous
Var8	11624	C	T	PP1ab - NSP9	synonymous

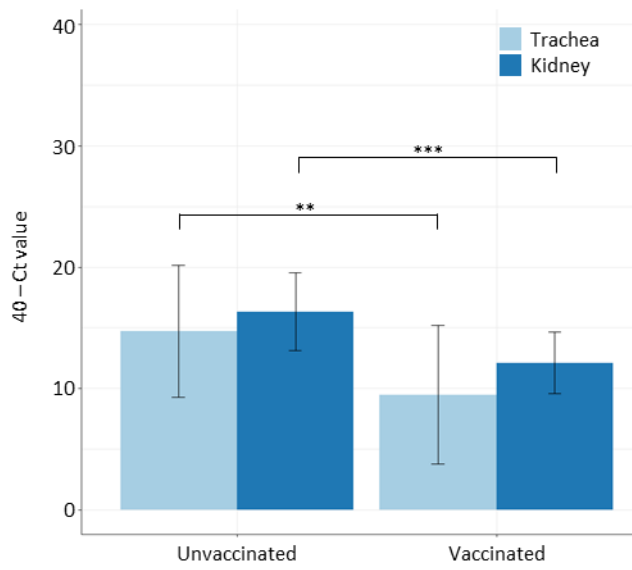


Figure 1: Comparison of relative mean of viral RNA in both tracheas and kidneys of vaccinated and unvaccinated birds of H120-D388 trial. Error bars represent standard deviation, statistic test (comparison of mean of two independent samples) were performed using Wilcoxon test. ** p-value below 0.01 ; *** pvalue below 0.001

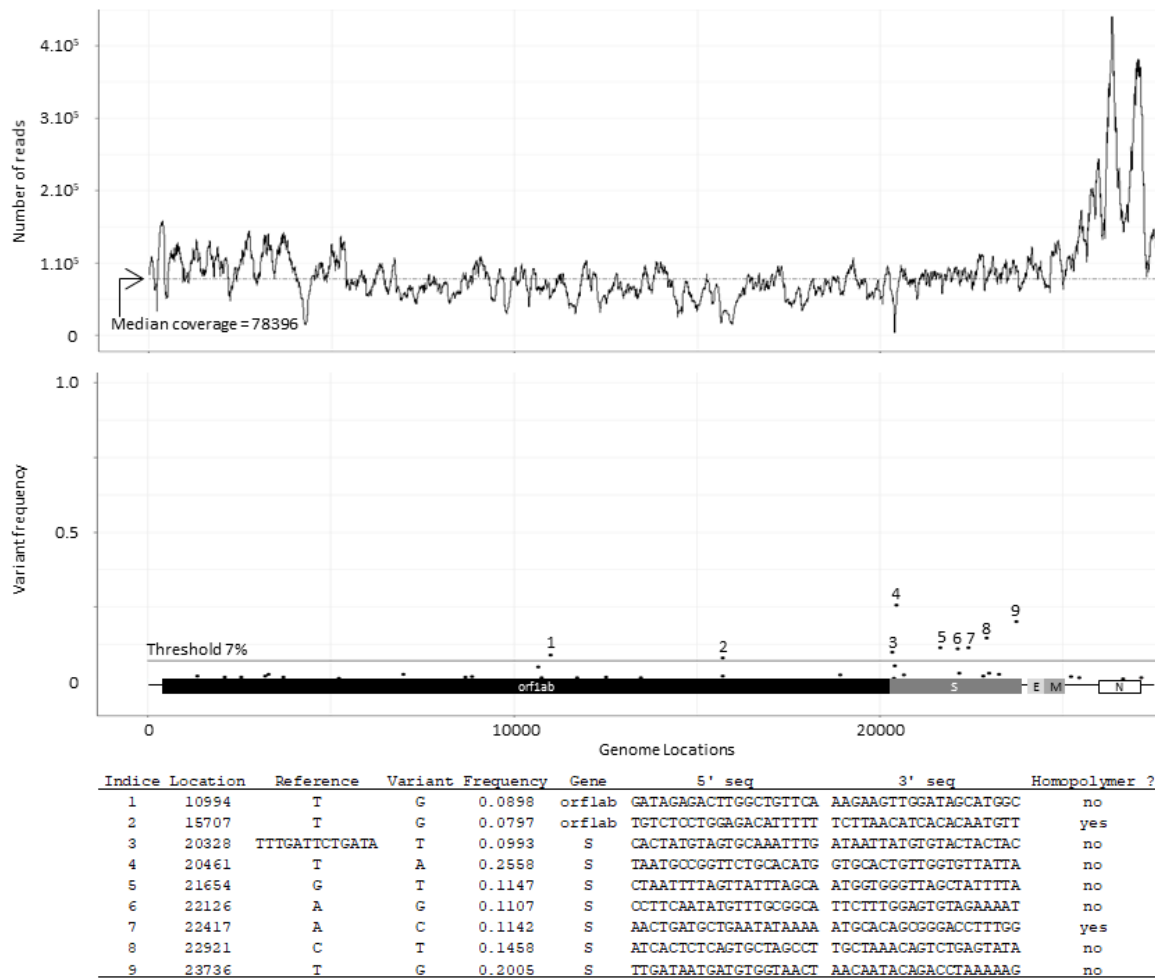


Figure 2: Genetic variant map of the D388 inoculum. Clarified virus stock of the D388 inoculum was deep-sequenced and NGS data were processed through VVV. Upper graphic shows genome coverage. Middle graphic presents detected genetic variants over the viral genome. Table give details of the detected genetic variants.

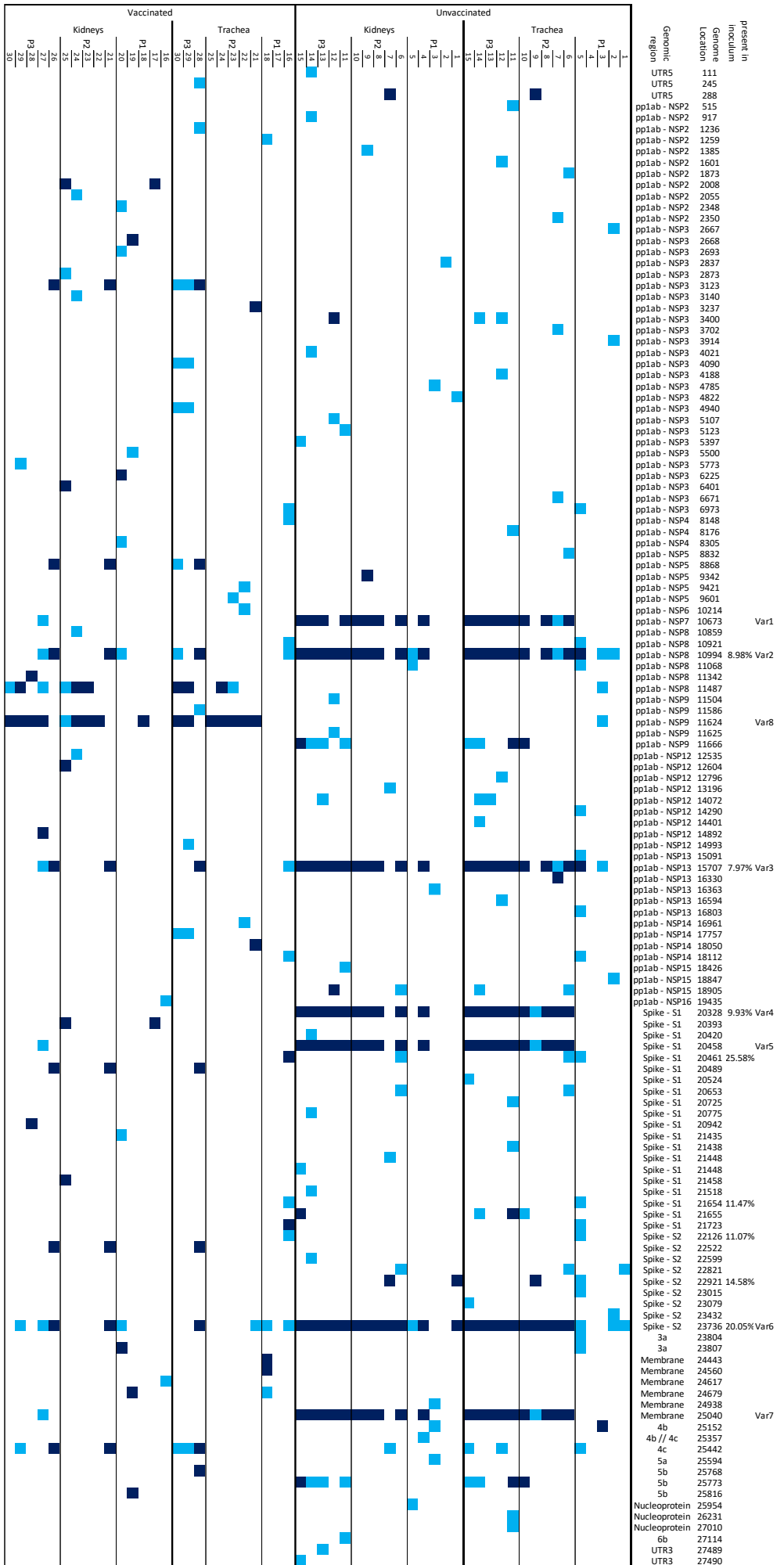


Figure 3: Evolution of D388 inoculum in trachea and kidneys of VB and NVB. D388 inoculum was passaged three times in VB and NVB and genome evolution was followed using Next Generation Sequencing (NGS). All NGS data from were processed through the VVV pipeline. Birds are presented in line, detected genetic variants are presented in column. Genomic regions together with genome locations are indicated above the graphic. Every time a variant was detected, a blue square was set: light blue colour if it was below the consensus, dark blue colour if it reached at consensus level. All other details about the detected genetic variants were voluntarily removed in order to lighten the figure. Complete details are given in Table S2.

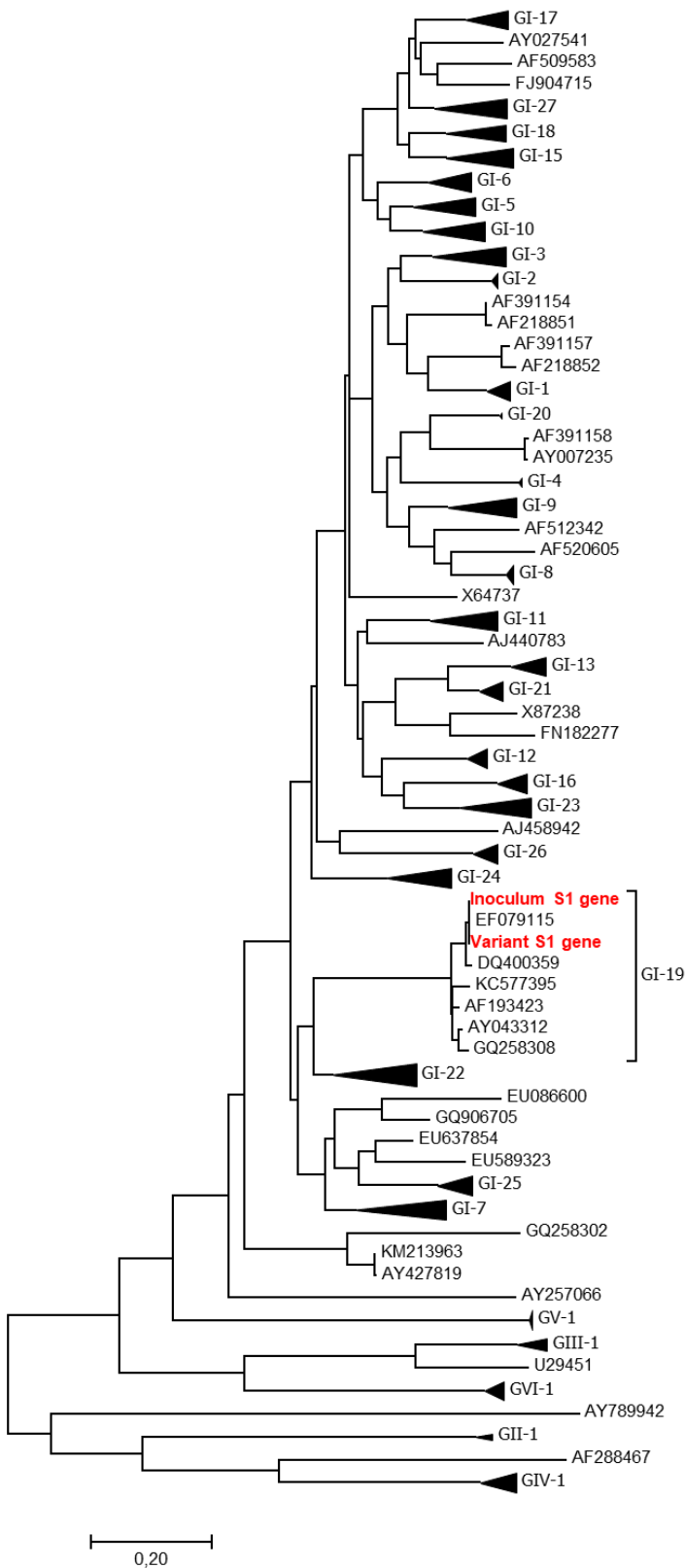


Figure 4: Phylogenetic analysis of the S1 protein of IBV. Phylogenetic tree was drawn using the Maximum Likelihood method based on the General Time Reversible model, a discrete Gamma distribution and the rate variation model that allowed for some sites to be evolutionarily invariable. The analysis involved 201 nucleotide sequences for a total of 1826 positions in the final dataset. Black triangles represent condensed group of sequence belonging to the same lineage. When not associated with a lineage, GenBank Accession Numbers represent unclassified sequences designed as Unique Variant in the Valastro classification (Valastro et al., 2016). “Inoculum S1 gene” is the S gene of the clarified D388 inoculum and “Variant S1 gene” is the S gene of the clarified 15-D388_{pass3Tr}NVB sample, both are indicated in red.

References

- Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402. <https://doi.org/10.1093/nar/25.17.3389>
- Callison, S.A., Hilt, D.A., Boynton, T.O., Sample, B.F., Robison, R., Swayne, D.E., Jackwood, M.W., 2006. Development and evaluation of a real-time Taqman RT-PCR assay for the detection of infectious bronchitis virus from infected chickens. *J. Virol. Methods* 138, 60–65. <https://doi.org/10.1016/j.jviromet.2006.07.018>
- Colvero, L.P., Villarreal, L.Y.B., Torres, C.A., Brandao, P.E., 2015. Assessing the economic burden of avian infectious bronchitis on poultry farms in Brazil: -EN- -FR- Évaluation de l'impact économique de la bronchite infectieuse aviaire dans les élevages de volailles brésiliens -ES- Determinación de la carga económica que impone la bronquitis infecciosa aviar a las explotaciones avícolas del Brasil. *Revue Scientifique et Technique de l'OIE* 34, 993–999. <https://doi.org/10.20506/rst.34.3.2411>
- Delaha, E.C., Reagan, R.L., Cook, S.R., Brueckner, A.L., 1954. Transmission of the avian infectious bronchitis virus to the cynomolgus monkey. *Cornell Vet* 44, 433–436.
- Flageul, A., Lucas, P., Hirchaud, E., Touzain, F., Blanchard, Y., Etteradossi, N., Brown, P., Grasland, B., 2021. Viral variant visualizer (VVV): A novel bioinformatic tool for rapid and simple visualization of viral genetic diversity. *Virus Research* 291, 198201. <https://doi.org/10.1016/j.virusres.2020.198201>
- Guy, J.S., 2015. Isolation and Propagation of Coronaviruses in Embryonated Eggs, in: Maier, H.J., Bickerton, E., Britton, P. (Eds.), *Coronaviruses: Methods and Protocols*, Methods in Molecular Biology. Springer, New York, NY, pp. 63–71. https://doi.org/10.1007/978-1-4939-2438-7_7
- Jackwood, M.W., 2012. Review of infectious bronchitis virus around the world. *Avian Dis.* 56, 634–641. <https://doi.org/10.1637/10227-043012-Review.1>
- Jackwood, M.W., Hall, D., Handel, A., 2012. Molecular evolution and emergence of avian gammacoronaviruses. *Infect. Genet. Evol.* 12, 1305–1311. <https://doi.org/10.1016/j.meegid.2012.05.003>
- Keep, S., Bickerton, E., Britton, P. (Eds.), 2015. Chapitre 11: Partial Purification of IBV and Subsequent Isolation of Viral RNA for Next-Generation Sequencing, in: *Coronaviruses: Methods and Protocols*, Methods in Molecular Biology. Humana Press.
- Kint, J., Maier, H.J., Jagt, E., 2015. Quantification of infectious bronchitis coronavirus by titration in vitro and in ovo. *Methods Mol. Biol.* 1282, 89–98. https://doi.org/10.1007/978-1-4939-2438-7_9
- Kumar, S., Stecher, G., Tamara, K., 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol* 33, 1870–1874. <https://doi.org/10.1093/molbev/msw054>
- Legnardi, M., Franzo, G., Koutoulis, K.C., Wiśniewski, M., Catelli, E., Tucciarone, C.M., Cecchinato, M., 2019. Vaccine or field strains: the jigsaw pattern of infectious bronchitis virus molecular epidemiology in Poland. *Poult. Sci.* <https://doi.org/10.3382/ps/pez473>
- Lontok, E., Corse, E., Machamer, C.E., 2004. Intracellular Targeting Signals Contribute to Localization of Coronavirus Spike Proteins near the Virus Assembly Site. *Journal of Virology* 78, 5913–5922. <https://doi.org/10.1128/JVI.78.11.5913-5922.2004>
- Ma, W., Goldberg, J., 2013. Rules for the recognition of dilysine retrieval motifs by coatomer. *The EMBO Journal* 32, 926–937. <https://doi.org/10.1038/emboj.2013.41>
- McKinley, E.T., Hilt, D.A., Jackwood, M.W., 2008. Avian coronavirus infectious bronchitis attenuated live vaccines undergo selection of subpopulations and mutations following vaccination. *Vaccine* 26, 1274–1284. <https://doi.org/10.1016/j.vaccine.2008.01.006>
- Nei, M., Kumar, S., 2000. *Molecular Evolution and Phylogenetics*. Oxford University Press.
- Reed, L.J., Muench, H., 1938. A SIMPLE METHOD OF ESTIMATING FIFTY PER CENT ENDPOINTS. *Am J Epidemiol* 27, 493–497. <https://doi.org/10.1093/oxfordjournals.aje.a118408>
- Schäffer, A.A., Aravind, L., Madden, T.L., Shavirin, S., Spouge, J.L., Wolf, Y.I., Koonin, E.V., Altschul, S.F., 2001. Improving the accuracy of PSI-BLAST protein database searches with composition-based statistics and other refinements. *Nucleic Acids Res.* 29, 2994–3005. <https://doi.org/10.1093/nar/29.14.2994>
- Toro, H., van Santen, V.L., Jackwood, M.W., 2012. Genetic diversity and selection regulates evolution of infectious bronchitis virus. *Avian Dis.* 56, 449–455. <https://doi.org/10.1637/10072-020212-Review.1>
- Valastro, V., Holmes, E.C., Britton, P., Fusaro, A., Jackwood, M.W., Cattoli, G., Monne, I., 2016. S1 gene-based phylogeny of infectious bronchitis virus: An attempt to harmonize virus classification. *Infect. Genet. Evol.* 39, 349–364. <https://doi.org/10.1016/j.meegid.2016.02.015>
- van Santen, V.L., Toro, H., 2008. Rapid selection in chickens of subpopulations within ArkDPI-derived infectious bronchitis virus vaccines. *Avian Pathol.* 37, 293–306. <https://doi.org/10.1080/03079450802043783>

Troisième partie
-
Discussion et Perspectives

9 Discussion Générale

L'objectif du travail présenté à travers ce manuscrit était d'étudier l'évolution de deux *Gammacoronavirus* aviaires, l'IBV et le TCoV, au cours de plusieurs passages des virus dans leurs hôtes en analysant les variants génétiques présents au sein des différentes populations virales avec ou sans pression de sélection pour l'IBV.

9.1 Améliorer la détection des variants génétiques

Dans un premier temps, un pipeline d'analyse a été développé afin d'automatiser la recherche de variants génétiques en utilisant un seul logiciel d'identification de variants, dit « *variant caller* ». Actuellement, bien que différents outils de recherches de variants existent (Koboldt et al., 2012; Lai et al., 2016; Wilm et al., 2012; Yang et al., 2013), aucun n'est capable de détecter l'ensemble des variants génétiques d'un échantillon (Sandmann et al., 2017). Ceci étant principalement à cause de la conception de ces outils qui utilisent des tests statistiques ou un ensemble de règles définies à respecter afin d'inclure ou d'exclure des variants génétiques potentiels. Pour pallier à ces problèmes, la combinaison de plusieurs outils de recherche de variants pourrait être envisagée afin d'augmenter la finesse de détection des variants génétiques sur la base d'une comparaison automatisée des résultats produits par chacun des outils de recherche de variants.

En pratique, un variant doit être présent en proportion suffisante pour être détectable. À titre d'exemple, avec un seuil de détection de 1 % et une profondeur de couverture de 1000 lectures à une position donnée, il doit y avoir un minimum de 10 lectures possédant le variant génétique pour que celui-ci soit détectable. Or si la profondeur de couverture diminue, le nombre minimal de lectures nécessaires pour qu'une séquence variante soit détectée diminue également et il devient donc difficile de distinguer un variant d'une erreur de séquençage ou de PCR. En définissant une profondeur de couverture médiane (500), avec un minimum (200) pour les positions où des variants génétiques sont détectés, et en considérant un seuil de détection supérieur à 7 %, ces erreurs de séquençages et de PCR ont pu être discriminées dans les études présentées dans ce manuscrit de thèse.

Afin de diminuer les erreurs de séquençage et de PCR et ainsi améliorer la détection de variants, il est également possible d'agir sur l'étape de préparation des bibliothèques pour le séquençage NGS. En effet, des protocoles particuliers ont été développés afin d'identifier et d'éliminer les erreurs pouvant survenir lors de l'amplification par PCR et du séquençage, permettant ainsi de réduire le taux d'erreur dans les données finales (Lu et al., 2020). Deux types de protocoles peuvent être distingués :

- le séquençage circulaire (Acevedo et al., 2014; Li et al., 2016; Lou et al., 2013). Cette méthode consiste à circulariser les chaînes d'acides nucléiques de sorte que le séquenceur puisse lire plusieurs fois la séquence d'une même chaîne circularisée. À la fin du processus, la lecture correspondra à une répétition d'une même chaîne d'acides nucléiques. Après traitement bioinformatique, ces répétitions seront fusionnées en une seule lecture dont la séquence correspondra à la majorité des nucléotides rencontrés à chaque position.
- le séquençage de chaînes d'acides nucléiques individuellement identifiées (Geller et al., 2016; Jabara et al., 2011). Cette méthode consiste à ajouter sur chaque chaîne d'acides nucléiques un identifiant unique, d'une taille définie et généré aléatoirement. Après séquençage, toutes les lectures possédant un même identifiant correspondront à une seule et unique chaîne nucléique de départ. Ces séquences pourront alors être comparées et fusionnées en une seule lecture dont la séquence correspondra à la majorité des nucléotides rencontrés à chaque position. Cette méthode présente également l'avantage d'identifier les lectures correspondant à des amplifications excessives de certain fragments (les duplicas de PCR).

Cependant, ces protocoles ne sont pas proposés par toutes les plateformes de séquençage haut débit, ce qui limite leur utilisation. Par ailleurs, même en utilisant ces protocoles, des erreurs de séquence peuvent apparaître au moment de la transcription inverse de l'ARN en ADNc, avec l'incorporation de nucléotides erronés par les transcriptases inverses.

9.2 L'évolution en l'absence d'immunité préexistante

Dans le cas de l'expérimentation avec IBV, il a pu être constaté qu'en absence d'immunité préexistante, l'évolution du virus IBV-D388 était très rapide. Après seulement trois passages du virus sur des poulets EOPS non vaccinés, sept changements de séquence consensus ont été observés sur l'ensemble du génome, dont quatre étaient le résultat d'une sélection de variants génétiques présents dans l'inoculum de départ. Ces mutations ont été retrouvées chez tous les sujets lors du troisième passage et présentes chez certains sujets dès le premier passage. Parmi elles, trois sont supposées induire d'importantes modifications de la séquence en acide aminé de la protéine de spicule.

Lors des passages successifs de la souche Fr-TCoV sur des dindes EOPS, cinq changements de séquence consensus ont pu être observés dès le premier passage du virus et se sont maintenus durant les 10 passages de l'expérience. Ces changements rapides de séquence consensus après seulement un passage sur dinde pourraient être attribués à une sélection de

variants génétiques présents dans l'inoculum de départ. Cependant en l'absence de données de séquençage de bonne qualité provenant de l'inoculum Fr-TCoV, il est impossible de confirmer ou d'infirmier cette hypothèse.

Ces différents résultats montrent que l'évolution des *Gammacoronavirus* aviaires est rapide lors de passages successifs des virus chez des animaux non vaccinés. Mais puisqu'à chaque passage, les conditions d'élevages des animaux étaient contrôlées, il est intéressant de se poser la question suivante : qu'est-ce qui a permis la sélection rapide de ces différents variants ?

Cette sélection rapide pourrait être attribuée à la nature des inocula IBV-D388 et Fr-TCoV qui sont des isolats viraux ayant été amplifiés respectivement quatre et six fois sur œufs embryonnés. Ceci fait écho à deux autres études, dans lesquelles la présence de variants génétiques au sein de plusieurs vaccins vivant IBV-Ark avait été identifiée par séquençage Sanger (McKinley et al., 2008; van Santen and Toro, 2008). Ces vaccins étaient le résultat de 50 passages sur œufs embryonnés et après inoculation chez des poulets, certains des variants composant la population virale du vaccin avaient été sélectionnés, entraînant des modifications différentes de séquence du gène S selon le vaccin (McKinley et al., 2008; van Santen and Toro, 2008). Bien que le nombre de passages sur œufs diffère grandement entre les souches virales (IBV-D388 et Fr-TCoV) et le virus vaccinal (IBV-Ark), il est possible qu'au cours des multiples passages sur œufs, des variants génétiques spécifiques de la culture sur œuf soient apparus et aient été sélectionnés, donnant naissance à une population virale parfaitement adaptée à ce système de culture. Puis, après inoculation de ces virus à leurs hôtes respectifs, une sélection rapide de variants génétiques ait été effectuée dès le premier passage pour s'adapter à l'hôte et permettre une meilleure réplication du virus dans l'hôte.

Au cours d'une étude dans laquelle ont été réalisées quatre séries de 106 passages sur œufs à partir d'une seule souche d'IBV-M41_CK, le séquençage des virus issus du 106^{ème} passage n'a révélé aucune modification du génome viral commune aux quatre amplifications (Oade et al., 2019). Cependant, cette conclusion se base sur des quantités de données NGS insuffisantes pour certains échantillons, se traduisant par de faibles profondeurs de couverture sur de nombreuses régions du génome qui peuvent empêcher la détection de variants et donc fausser les résultats.

Par ailleurs, l'étude d'Oade et al. (2019) est contredite pas des travaux antérieurs qui montrent qu'il existe des séquences virales spécifiques de la réplifications d'IBV sur œuf, et d'autres spécifiques de la réplification d'IBV sur poulet (Callison et al., 2001; Cavanagh et al., 2005; McKinley et al., 2008). Au cours d'expériences alternant réplification sur poulet et réplification sur œuf, il a été montré qu'une souche IBV-793b issue du poulet avait changé de séquence consensus au niveau du gène S (le seul à avoir été séquencé) après un seul passage sur œuf. Dans ce cas

précis, la souche issue du poulet possédait un T en position 283 du gène S (différente des positions analysées lors de l'expérimentation IBV-D388) et après un passage sur œuf un G est devenu le nucléotide majoritaire à cette position, mais un T était également retrouvé en tant que variant génétique (Callison et al., 2001; Cavanagh et al., 2005). Puis, lorsque le virus issu de l'œuf a été inoculé à un poulet, le T était de nouveau majoritaire. Après 10 passages de ce virus sur poulet le T était toujours majoritaire mais le G n'était plus détecté. Enfin, après l'inoculation à l'œuf du virus issu du 10^{ème} passage sur poulet, le G était de nouveau majoritaire et le T n'était plus détecté (Cavanagh et al., 2005). Une autre étude, utilisant une souche vaccinale IBV-Ark (produite sur œuf) a révélé l'apparition de mutations du génome vaccinal après un seul passage sur poulets. Ensuite, le virus issu du poulet a été utilisé pour une série de 10 passages sur œufs, au cours de laquelle certaines des mutations apparues chez le poulet ont été modifiées pour de nouveau correspondre à la séquence du génome vaccinal (McKinley et al., 2008).

Ces changements de séquence consensus observés lors des expérimentations IBV et TCoV sont donc possiblement induits par le changement de système de répllication du virus (œuf/poulet). Si tel est le cas, il pourrait donc être imaginé qu'aucun autre changement « rapide » et durable ne survient tant que le virus reste dans le même hôte et dans les mêmes conditions de répllication. Par ailleurs, lors de l'expérimentation TCoV, les mêmes changements « rapides » et durables de la séquence consensus ont pu être observés dans les trois répliques, indiquant donc une répétabilité des changements génétiques. En conséquence, tant que le virus se réplique dans l'hôte pour lequel il est adapté, seuls des changements occasionnels et aléatoires pourraient survenir.

9.3 L'évolution en présence d'une immunité préexistante

Au début des années 2000, il a été suggéré que l'émergence de la souche IBV-GA98 était le résultat de la mise en place d'une vaccination contre la souche circulante de l'époque IBV-DE072 (Lee et al., 2001). Par ailleurs, une autre étude suggérait une évolution différente d'une souche vaccinale IBV-Ark si les sujets avaient un système immunitaire fonctionnel ou déficient, résultat d'une infection par des virus immunosuppresseurs (tels que le virus de la bursite infectieuse ou le virus de l'anémie du poulet) (Gallardo et al., 2012). De plus, une autre étude a montré qu'une souche terrain IBV-Ark administrée à des poulets de 21 jours d'âge, préalablement vaccinés ou non à 5 jours d'âge avec un vaccin vivant IBV-Ark, évoluait différemment si les sujets avaient été vaccinés ou non (Toro et al., 2012). Cependant, ce résultat doit être considéré avec précaution, car au moment de l'inoculation avec un virus épreuve IBV-Ark (à 21 jours d'âge, soit 16 jours post vaccination), aucun contrôle n'a été effectué afin de vérifier l'absence d'excrétion de virus vaccinal par les sujets vaccinés au moment de l'épreuve, alors même qu'un travail précédent

montrait que la souche vaccinale IBV-Ark pouvait être séquencée chez certains sujets, vingt jours après l'inoculation du vaccin (Gallardo et al., 2010). Or il se trouve que les populations virales identifiées après épreuve avec la souche terrain, correspondent à celles identifiées lors d'études analysant l'évolution de la souche vaccinale IBV-Ark (Gallardo et al., 2010; van Santen and Toro, 2008), ce qui tend à conclure que les populations virales identifiées dans l'étude de Toro et al. (2012) sont issues du vaccin et non de la souche épreuve. Dans de telles circonstances, il est logique que les populations virales identifiées chez les sujets vaccinés et non vaccinés soient différentes puisqu'elles ne sont pas issues du même inoculum.

Lors de l'expérimentation IBV-D388, une sélection de variants génétiques a été observée chez les sujets vaccinés avec IBV-H120, indiquant qu'une évolution virale est possible lorsque l'immunité vaccinale ne suffit pas à stopper la réplication du virus. Ceci est le cas sur le terrain, lorsque le virus vaccinal ne correspond pas aux souches circulantes (Colvero et al., 2015), mais pourrait également être le cas, lorsque des doses insuffisantes de vaccins sont administrées à des sujets, entraînant une mauvaise stimulation de leur système immunitaire. Par ailleurs, la sélection de variants génétiques qui s'est effectuée chez les sujets vaccinés est différente de celle observée chez les sujets non vaccinés, mais certaines des variations génétiques sélectionnées chez les sujets non vaccinés, ont également été détectées chez les sujets vaccinés sans qu'aucune sélection ne s'opère. Devant de tels résultats, il semble donc que la présence d'une immunité préexistante soit capable de modifier l'évolution naturelle de l'IBV, en empêchant la survenue de mutations qui seraient apparues en absence d'immunité préexistante, et en permettant l'apparition de mutations qui ne seraient pas survenues en absence d'immunité. Ces résultats mettent donc évidence la nécessité d'utiliser un vaccin capable de stopper efficacement la réplication du virus.

10 Perspectives

Sur la base des expériences réalisées dans ce projet de thèse, de nouvelles analyses pourraient être effectuées.

Dans le cadre de l'expérimentation IBV, il serait intéressant d'effectuer le séquençage des différents passages de la souche IBV-M41 chez des animaux non vaccinés afin de comparer l'évolution avec celle observée pour IBV-D388. En effet, IBV-M41 étant également un isolat passé plusieurs fois sur œufs, il serait intéressant de voir si une sélection rapide de variants génétiques est aussi observée.

Ensuite, recommencer l'expérimentation IBV décrite dans ce manuscrit, en utilisant les mêmes stocks viraux de départ, permettrait de comparer les résultats avec la première expérience. En cas de résultats identiques, cela confirmerait que l'évolution observée est soumise à des phénomènes spécifiques qui restent encore à identifier. En cas de résultats différents, cela indiquerait que les résultats observés pourraient être dus au hasard. Cependant, cette dernière hypothèse semble peu probable car les résultats obtenus lors de l'expérimentation TCoV montrent que les sélections au niveau de certaines positions du génome ont été détectés dans les trois répliques réalisées avec le TCoV. Il serait intéressant d'étudier l'impact des modifications génétiques sélectionnées lors de l'expérimentation IBV vis-à-vis d'un échappement à la réponse immunitaire induite contre l'inoculum IBV-D388 de départ. Ainsi, les virus variants seraient utilisés lors d'une nouvelle expérimentation animale dans laquelle les sujets seraient vaccinés contre l'inoculum D388 et éprouvés avec les virus variants.

Il pourrait être également intéressant d'étudier l'impact du type d'immunité dans l'évolution virale, en se posant la question suivante : les vaccins inactivés modifient-ils l'évolution virale ? Pour cela, l'expérimentation IBV pourrait être de nouveau réalisée en ajoutant un troisième groupe de sujets recevant un vaccin inactivé.

Dans le cas de l'expérimentation TCoV, séquencer le contenu des caeca de dindes lors de chacun des 10 passages pourrait permettre d'améliorer le suivi de la dynamique évolutive des populations virales et ainsi confirmer si les variants sélectionnés ont été maintenus au cours des 10 passages, ou si leur détection était fluctuante. Par ailleurs, l'intégralité des intestins de dindes a été récupéré au cours de cette expérience, permettant ainsi d'évaluer l'évolution du TCoV dans les différents segments intestinaux et de comparer avec l'évolution observée dans les caeca.

Les travaux qui ont été menés au cours de ce projet de thèse ont permis d'apporter des éléments de compréhension quant à l'évolution de deux *Gammacoronavirus* aviaires mais ont également soulevé de nombreuses questions. Il a été constaté que l'évolution du virus IBV était possible et différente de l'évolution naturelle, lorsque le virus se réplique en présence d'une

immunité préexistante non protectrice, induite par l'usage d'un vaccin IBV non adapté à la souche épreuve.

Or, il se trouve que la vaccination anti-IBV est une problématique complexe de la filière avicole. En effet, de nombreux vaccins anti-IBV existent et beaucoup sont des vaccins vivants atténués après de multiples passages sur œufs, qui peuvent donc subir un retour à la virulence après passage sur poulet, par sélection de variants présents dans les populations virales vaccinales (comme cela a pu être observé chez les animaux non vaccinés de l'expérimentation IBV-D388) ou par recombinaison avec des souches virales circulantes, ou d'autres souches vaccinales lors de l'administrations simultanées ou différées de plusieurs vaccins vivants dans le but d'augmenter les protections croisées. Devant ces risques, il est nécessaire de travailler sur le développement de nouveaux vaccins capables d'allier efficacité de protection et sécurité d'emploi.

11 Références bibliographiques

- Acevedo, A., Brodsky, L., Andino, R., 2014. Mutational and fitness landscapes of an RNA virus revealed through population sequencing. *Nature* 505, 686–690. <https://doi.org/10.1038/nature12861>
- Adzhar, A.B., Shaw, K., Britton, P., Cavanagh, D., 1995. Avian infectious bronchitis virus: differences between 793/B and other strains. *Vet. Rec.* 136, 548. <https://doi.org/10.1136/vr.136.21.548-a>
- Alonso-Caplen, F.V., Matsuoka, Y., Wilcox, G.E., Compans, R.W., 1984. Replication and morphogenesis of avian coronavirus in Vero cells and their inhibition by monensin. *Virus Res.* 1, 153–167. [https://doi.org/10.1016/0168-1702\(84\)90070-4](https://doi.org/10.1016/0168-1702(84)90070-4)
- Animas, S.B., Otsuki, K., Hanayama, M., Sanekata, T., Tsubokura, M., 1994a. Experimental infection with avian infectious bronchitis virus (Kagoshima-34 strain) in chicks at different ages. *J. Vet. Med. Sci.* 56, 443–447. <https://doi.org/10.1292/jvms.56.443>
- Animas, S.B., Otsuki, K., Tsubokura, M., Cook, J.K., 1994b. Comparison of the susceptibility of chicks of different ages to infection with nephrosis/nephritis-causing strain of infectious bronchitis virus. *J. Vet. Med. Sci.* 56, 449–453. <https://doi.org/10.1292/jvms.56.449>
- Aouadi, W., Eydoux, C., Coutard, B., Martin, B., Debart, F., Vasseur, J.J., Contreras, J.M., Morice, C., Quérat, G., Jung, M.-L., Canard, B., Guillemot, J.-C., Decroly, E., 2017. Toward the identification of viral cap-methyltransferase inhibitors by fluorescence screening assay. *Antiviral Research* 144, 330–339. <https://doi.org/10.1016/j.antiviral.2017.06.021>
- Awe, O.O., Kang, K., Ibrahim, M., Ali, A., Elaish, M., Saif, Y.M., Lee, C.-W., 2015. Age-Related Susceptibility of Turkeys to Enteric Viruses. *Avian Dis.* 59, 207–212. <https://doi.org/10.1637/10907-071514-Reg>
- Bande, F., Arshad, S.S., Omar, A.R., Hair-Bejo, M., Mahmuda, A., Nair, V., 2017. Global distributions and strain diversity of avian infectious bronchitis virus: a review. *Anim Health Res Rev* 18, 70–83. <https://doi.org/10.1017/S1466252317000044>
- Barnes, H.J., Guy, J.S., Vaillancourt, J.P., 2000. Poultry enteritis complex. *Rev. - Off. Int. Epizoot.* 19, 565–588. <https://doi.org/10.20506/rst.19.2.1234>
- Beaudette, F.R., Hudson, C.B., 1937. Cultivation Of The Virus Of Infectious Bronchitis 90, 51–60.
- Benyeda, Z., Mató, T., Süveges, T., Szabó, E., Kardi, V., Abonyi-Tóth, Z., Rusvai, M., Palya, V., 2009. Comparison of the pathogenicity of QX-like, M41 and 793/B infectious bronchitis strains from different pathological conditions. *Avian Pathol.* 38, 449–456. <https://doi.org/10.1080/03079450903349196>
- Benyeda, Z., Szeredi, L., Mató, T., Süveges, T., Balka, G., Abonyi-Tóth, Z., Rusvai, M., Palya, V., 2010. Comparative Histopathology and Immunohistochemistry of QX-like, Massachusetts and 793/B Serotypes of Infectious Bronchitis Virus Infection in Chickens. *J Comp Pathol* 143, 276–283. <https://doi.org/10.1016/j.jcpa.2010.04.007>
- Berger, R., 2017. Coronavirus entérique de la pintade en France : pathogénèse, transmission et évolution (other).
- Berry, D.M., Cruickshank, J.G., Chu, H.P., Wells, R.J.H., 1964. The structure of infectious bronchitis virus. *Virology* 23, 403–407. [https://doi.org/10.1016/0042-6822\(64\)90263-6](https://doi.org/10.1016/0042-6822(64)90263-6)
- Bhattacharjee, P.S., Jones, R.C., 1997. Susceptibility of organ cultures from chicken tissues for strains of infectious bronchitis virus isolated from the intestine. *Avian Pathol.* 26, 553–563. <https://doi.org/10.1080/03079459708419234>
- Boroomand, Z., Asasi, K., Mohammadi, A., 2012. Pathogenesis and tissue distribution of avian infectious bronchitis virus isolate IRFIBV32 (793/B serotype) in experimentally infected broiler chickens. *ScientificWorldJournal* 2012, 402537. <https://doi.org/10.1100/2012/402537>

- Boroomand, Z., PhD, D., (DVM, R., Asasi, K., Mohammadi, A., (DVM, A., Hosseini, S.M., 2011. Isolation and identification of a new isolate of avian infectious bronchitis virus IRFIBV32 and a study of its pathogenicity 15, 366–380.
- Bosch, B.J., van der Zee, R., de Haan, C.A.M., Rottier, P.J.M., 2003. The coronavirus spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex. *J. Virol.* 77, 8801–8811. <https://doi.org/10.1128/jvi.77.16.8801-8811.2003>
- Boursnell, M.E., Brown, T.D., Foulds, I.J., Green, P.F., Tomley, F.M., Binns, M.M., 1987. Completion of the sequence of the genome of the coronavirus avian infectious bronchitis virus. *J. Gen. Virol.* 68 (Pt 1), 57–77. <https://doi.org/10.1099/0022-1317-68-1-57>
- Bouvet, M., Lugari, A., Posthuma, C.C., Zevenhoven, J.C., Bernard, S., Betzi, S., Imbert, I., Canard, B., Guillemot, J.-C., Lécine, P., Pfefferle, S., Drosten, C., Snijder, E.J., Decroly, E., Morelli, X., 2014. Coronavirus Nsp10, a Critical Co-factor for Activation of Multiple Replicative Enzymes. *Journal of Biological Chemistry* 289, 25783–25796. <https://doi.org/10.1074/jbc.M114.577353>
- Bouwman, K.M., Habraeken, N., Laconi, A., Berends, A.J., Groenewoud, L., Alders, M., Kemp, V., Verheije, M.H., 2020. N-glycosylation of infectious bronchitis virus M41 spike determines receptor specificity. *Journal of General Virology.* <https://doi.org/10.1099/jgv.0.001408>
- Bouwman, Parsons, L.M., Berends, A.J., de Vries, R.P., Cipollo, J.F., Verheije, M.H., 2020. Three Amino Acid Changes in Avian Coronavirus Spike Protein Allow Binding to Kidney Tissue. *J. Virol.* 94. <https://doi.org/10.1128/JVI.01363-19>
- Brierley, I., Boursnell, M.E., Binns, M.M., Bilimoria, B., Blok, V.C., Brown, T.D., Inglis, S.C., 1987. An efficient ribosomal frame-shifting signal in the polymerase-encoding region of the coronavirus IBV. *EMBO J.* 6, 3779–3785.
- Brierley, I., Jenner, A.J., Inglis, S.C., 1992. Mutational analysis of the “slippery-sequence” component of a coronavirus ribosomal frameshifting signal. *Journal of Molecular Biology* 227, 463–479. [https://doi.org/10.1016/0022-2836\(92\)90901-U](https://doi.org/10.1016/0022-2836(92)90901-U)
- Brierley, I., Rolley, N.J., Jenner, A.J., Inglis, S.C., 1991. Mutational analysis of the RNA pseudoknot component of a coronavirus ribosomal frameshifting signal. *J. Mol. Biol.* 220, 889–902.
- Brown, P.A., Courtillon, C., Weerts, E.A.W.S., Andraud, M., Allée, C., Vendembeuche, A., Amelot, M., Rose, N., Verheije, M.H., Eterradossi, N., 2018. Transmission Kinetics and histopathology induced by European Turkey Coronavirus during experimental infection of specific pathogen free turkeys. *Transbound Emerg Dis* 66, 234–242. <https://doi.org/10.1111/tbed.13006>
- Brown, P.A., Touzain, F., Briand, F.X., Gouilh, A.M., Courtillon, C., Allée, C., Lemaitre, E., De Boisséson, C., Blanchard, Y., Eterradossi, N., 2016. First complete genome sequence of European turkey coronavirus suggests complex recombination history related with US turkey and guinea fowl coronaviruses. *J. Gen. Virol.* 97, 110–120. <https://doi.org/10.1099/jgv.0.000338>
- Bushnell, L.D., Brandly, C.A., 1933. Laryngotracheitis in Chicks*. *Poultry Science* 12, 55–60. <https://doi.org/10.3382/ps.0120055>
- Callison, S.A., Hilt, D.A., Boynton, T.O., Sample, B.F., Robison, R., Swayne, D.E., Jackwood, M.W., 2006. Development and evaluation of a real-time Taqman RT-PCR assay for the detection of infectious bronchitis virus from infected chickens. *J. Virol. Methods* 138, 60–65. <https://doi.org/10.1016/j.jviromet.2006.07.018>
- Callison, S.A., Jackwood, M.W., Hilt, D.A., 2001. Molecular characterization of infectious bronchitis virus isolates foreign to the United States and comparison with United States isolates. *Avian Dis.* 45, 492–499.

- Cao, J., Wu, C.-C., Lin, T.L., 2008a. Turkey coronavirus non-structure protein NSP15--an endoribonuclease. *Intervirology* 51, 342–351. <https://doi.org/10.1159/000175837>
- Cao, J., Wu, C.-C., Lin, T.L., 2008b. Complete nucleotide sequence of polyprotein gene 1 and genome organization of turkey coronavirus. *Virus Res* 136, 43–49. <https://doi.org/10.1016/j.virusres.2008.04.015>
- Cavanagh, D., 2005. Coronaviruses in poultry and other birds. *Avian Pathol.* 34, 439–448. <https://doi.org/10.1080/03079450500367682>
- Cavanagh, D., 2001. A nomenclature for avian coronavirus isolates and the question of species status. *Avian Pathol.* 30, 109–115. <https://doi.org/10.1080/03079450120044506>
- Cavanagh, D., Davis, P.J., Mockett, A.P., 1988. Amino acids within hypervariable region 1 of avian coronavirus IBV (Massachusetts serotype) spike glycoprotein are associated with neutralization epitopes. *Virus Res.* 11, 141–150. [https://doi.org/10.1016/0168-1702\(88\)90039-1](https://doi.org/10.1016/0168-1702(88)90039-1)
- Cavanagh, D., Davis, P.J., Pappin, D.J., Binns, M.M., Bournsnel, M.E., Brown, T.D., 1986. Coronavirus IBV: partial amino terminal sequencing of spike polypeptide S2 identifies the sequence Arg-Arg-Phe-Arg-Arg at the cleavage site of the spike precursor polypeptide of IBV strains Beaudette and M41. *Virus Res.* 4, 133–143. [https://doi.org/10.1016/0168-1702\(86\)90037-7](https://doi.org/10.1016/0168-1702(86)90037-7)
- Cavanagh, D., Mawditt, K., Sharma, M., Drury, S.E., Ainsworth, H.L., Britton, P., Gough, R.E., 2001. Detection of a coronavirus from turkey poult in Europe genetically related to infectious bronchitis virus of chickens. *Avian Pathol.* 30, 355–368. <https://doi.org/10.1080/03079450120066368>
- Cavanagh, D., Mawditt, K., Welchman, D.D.B., Britton, P., Gough, R.E., 2002. Coronaviruses from pheasants (*Phasianus colchicus*) are genetically closely related to coronaviruses of domestic fowl (infectious bronchitis virus) and turkeys. *Avian Pathol.* 31, 81–93. <https://doi.org/10.1080/03079450120106651>
- Cavanagh, D., Picault, J.-P., Gough, R., Hess, M., Mawditt, K., Britton, P., 2005. Variation in the spike protein of the 793/B type of infectious bronchitis virus, in the field and during alternate passage in chickens and embryonated eggs. *Avian Pathol.* 34, 20–25. <https://doi.org/10.1080/03079450400025414>
- Cheng, J., Huo, C., Zhao, J., Liu, T., Li, X., Yan, S., Wang, Z., Hu, Y., Zhang, G., 2018. Pathogenicity differences between QX-like and Mass-type infectious bronchitis viruses. *Veterinary Microbiology* 213, 129–135. <https://doi.org/10.1016/j.vetmic.2017.11.027>
- Cheng, J., Zhao, Y., Xu, G., Zhang, K., Jia, W., Sun, Y., Zhao, J., Xue, J., Hu, Y., Zhang, G., 2019. The S2 Subunit of QX-type Infectious Bronchitis Coronavirus Spike Protein Is an Essential Determinant of Neurotropism. *Viruses* 11. <https://doi.org/10.3390/v11100972>
- Chu, V.C., McElroy, L.J., Aronson, J.M., Oura, T.J., Harbison, C.E., Bauman, B.E., Whittaker, G.R., 2007. Feline aminopeptidase N is not a functional receptor for avian infectious bronchitis virus. *Virology* 4, 20. <https://doi.org/10.1186/1743-422X-4-20>
- Chu, V.C., McElroy, L.J., Chu, V., Bauman, B.E., Whittaker, G.R., 2006. The avian coronavirus infectious bronchitis virus undergoes direct low-pH-dependent fusion activation during entry into host cells. *J. Virol.* 80, 3180–3188. <https://doi.org/10.1128/JVI.80.7.3180-3188.2006>
- Circella, E., Camarda, A., Martella, V., Bruni, G., Lavazza, A., Buonavoglia, C., 2007. Coronavirus associated with an enteric syndrome on a quail farm. *Avian Pathology* 36, 251–258. <https://doi.org/10.1080/03079450701344738>
- Colvero, L.P., Villarreal, L.Y.B., Torres, C.A., Brandao, P.E., 2015. Assessing the economic burden of avian infectious bronchitis on poultry farms in Brazil: -EN- -FR- Évaluation de l'impact économique de la bronchite infectieuse aviaire dans les élevages de volailles brésiliens -ES- Determinación de la carga económica que impone la bronquitis infecciosa

- aviar a las explotaciones avícolas del Brasil. *Revue Scientifique et Technique de l'OIE* 34, 993–999. <https://doi.org/10.20506/rst.34.3.2411>
- Cook, J.K.A., Jackwood, M., Jones, R.C., 2012. The long view: 40 years of infectious bronchitis research. *Avian Pathology* 41, 239–250. <https://doi.org/10.1080/03079457.2012.680432>
- Coria, M.F., Peterson, J.K., 1971. Adaptation and propagation of avian infectious bronchitis virus in embryonic turkey kidney cell cultures. *Avian Dis.* 15, 22–27.
- Corse, E., Machamer, C.E., 2000. Infectious bronchitis virus E protein is targeted to the Golgi complex and directs release of virus-like particles. *J. Virol.* 74, 4319–4326. <https://doi.org/10.1128/jvi.74.9.4319-4326.2000>
- Cottam, E.M., Maier, H.J., Manifava, M., Vaux, L.C., Chandra-Schoenfelder, P., Gerner, W., Britton, P., Ktistakis, N.T., Wileman, T., 2011. Coronavirus nsp6 proteins generate autophagosomes from the endoplasmic reticulum via an omegasome intermediate. *Autophagy* 7, 1335–1347. <https://doi.org/10.4161/auto.7.11.16642>
- Cottam, E.M., Whelband, M.C., Wileman, T., 2014. Coronavirus NSP6 restricts autophagosome expansion. *Autophagy* 10, 1426–1441. <https://doi.org/10.4161/auto.29309>
- Coutard, B., Valle, C., de Lamballerie, X., Canard, B., Seidah, N.G., Decroly, E., 2020. The spike glycoprotein of the new coronavirus 2019-nCoV contains a furin-like cleavage site absent in CoV of the same clade. *Antiviral Res* 176, 104742. <https://doi.org/10.1016/j.antiviral.2020.104742>
- Cunningham, C.H., Spring, M.P., 1965. SOME STUDIES OF INFECTIOUS BRONCHITIS VIRUS IN CELL CULTURE. *Avian Dis.* 9, 182–192.
- Cunningham, C.H., Spring, M.P., Nazerian, K., 1972. Replication of avian infectious bronchitis virus in African green monkey kidney cell line VERO. *J. Gen. Virol.* 16, 423–427. <https://doi.org/10.1099/0022-1317-16-3-423>
- Cunningham, C.H., Stuart, H.O., 1947. Cultivation of the virus of infectious bronchitis of chickens in embryonated chicken eggs. *Am. J. Vet. Res.* 8, 209–212.
- Dalton, K., Casais, R., Shaw, K., Stirrups, K., Evans, S., Britton, P., Brown, T.D.K., Cavanagh, D., 2001. cis-Acting Sequences Required for Coronavirus Infectious Bronchitis Virus Defective-RNA Replication and Packaging. *J Virol* 75, 125–133. <https://doi.org/10.1128/JVI.75.1.125-133.2001>
- Darbyshire, J.H., 1978. Organ culture in avian virology: A review. *Avian Pathology* 7, 321–335. <https://doi.org/10.1080/03079457808418287>
- Darbyshire, J.H., Cook, J.K., Peters, R.W., 1976. Organ culture studies on the efficiency of infection of chicken tissues with avian infectious bronchitis virus. *Br J Exp Pathol* 57, 443–454.
- Davies, H.A., Dourmashkin, R.R., Macnaughton, M.R., 1981. Ribonucleoprotein of avian infectious bronchitis virus. *J. Gen. Virol.* 53, 67–74. <https://doi.org/10.1099/0022-1317-53-1-67>
- De Wit, J.J.S., Ter Veen, C., Koopman, H.C.R., 2020. Effect of IBV D1466 on egg production and egg quality and the effect of heterologous priming to increase the efficacy of an inactivated IBV vaccine. *Avian Pathol.* 49, 185–192. <https://doi.org/10.1080/03079457.2019.1710462>
- De Wit, S.J.J., Cook, J.K.A., van der Heijden, H.M.J.F., 2011a. Infectious bronchitis virus variants: a review of the history, current situation and control measures. *Avian Pathol.* 40, 223–235. <https://doi.org/10.1080/03079457.2011.566260>
- De Wit, S.J.J., de Jong, M.C., Pijpers, A., Verheijden, J.H., 1998. Transmission of infectious bronchitis virus within vaccinated and unvaccinated groups of chickens. *Avian Pathol.* 27, 464–471. <https://doi.org/10.1080/03079459808419370>
- De Wit, S.J.J., Nieuwenhuisen-van Wilgen, J., Hoogkamer, A., van de Sande, H., Zuidam, G.J., Fabri, T.H.F., 2011b. Induction of cystic oviducts and protection against early challenge

- with infectious bronchitis virus serotype D388 (genotype QX) by maternally derived antibodies and by early vaccination. *Avian Pathol.* 40, 463–471. <https://doi.org/10.1080/03079457.2011.599060>
- Dea, S., Garzon, S., Strykowski, H., Tijssen, P., 1989a. Ultrastructure and protein A-gold immunolabelling of HRT-18 cells infected with turkey enteric coronavirus. *Veterinary Microbiology* 20, 21–33. [https://doi.org/10.1016/0378-1135\(89\)90004-7](https://doi.org/10.1016/0378-1135(89)90004-7)
- Dea, S., Garzon, S., Tijssen, P., 1989b. Identification and location of the structural glycoproteins of a tissue culture-adapted turkey enteric coronavirus. *Arch. Virol.* 106, 221–237. <https://doi.org/10.1007/bf01313955>
- Dea, S., Garzon, S., Tijssen, P., 1989c. Isolation and trypsin-enhanced propagation of turkey enteric (bluecomb) coronaviruses in a continuous human rectal adenocarcinoma cell line. *American Journal of Veterinary Research* 50, 1310–1318.
- Dea, S., Tijssen, P., 1988a. Identification of the structural proteins of turkey enteric coronavirus. *Arch Virol* 99, 173–186. <https://doi.org/10.1007/BF01311068>
- Dea, S., Tijssen, P., 1988b. Viral agents associated with outbreaks of diarrhea in turkey flocks in Quebec. *Can J Vet Res* 52, 53–57.
- Dea, S., Verbeek, A., Tijssen, P., 1991. Transmissible enteritis of turkeys: experimental inoculation studies with tissue-culture-adapted turkey and bovine coronaviruses. *Avian Diseases* 35, 767–777.
- Dea, S., Verbeek, A.J., Tijssen, P., 1990. Antigenic and genomic relationships among turkey and bovine enteric coronaviruses. *Journal of Virology* 64, 3112–3118. <https://doi.org/10.1128/JVI.64.6.3112-3118.1990>
- Decroly, E., Imbert, I., Coutard, B., Bouvet, M., Selisko, B., Alvarez, K., Gorbalenya, A.E., Snijder, E.J., Canard, B., 2008. Coronavirus Nonstructural Protein 16 Is a Cap-0 Binding Enzyme Possessing (Nucleoside-2'O)-Methyltransferase Activity. *Journal of Virology* 82, 8071–8084. <https://doi.org/10.1128/JVI.00407-08>
- Doherty, P.C., 1967. Titration of avian infectious bronchitis virus in the tissues of experimentally infected chickens. *Aust. Vet. J.* 43, 575–578. <https://doi.org/10.1111/j.1751-0813.1967.tb04805.x>
- Domanska-Blicharz, K., Seroka, A., Lisowska, A., Tomczyk, G., Minta, Z., 2010. TURKEY CORONAVIRUS IN POLAND - PRELIMINARY RESULTS. *Bull Vet Inst Pulawy* 473–477.
- Domingo, E., Sabo, D., Taniguchi, T., Weissmann, C., 1978. Nucleotide sequence heterogeneity of an RNA phage population. *Cell* 13, 735–744. [https://doi.org/10.1016/0092-8674\(78\)90223-4](https://doi.org/10.1016/0092-8674(78)90223-4)
- Doyle, N., Neuman, B.W., Simpson, J., Hawes, P.C., Mantell, J., Verkade, P., Alrashedi, H., Maier, H.J., 2018. Infectious Bronchitis Virus Nonstructural Protein 4 Alone Induces Membrane Pairing. *Viruses* 10. <https://doi.org/10.3390/v10090477>
- Dubose, R.T., Grumbles, L.C., Flowers, A.I., 1960. Differentiation of quail bronchitis virus and infectious bronchitis virus by heat stability. *Am. J. Vet. Res.* 21, 740–743.
- Ducatez, M.F., Liais, E., Croville, G., Guérin, J.-L., 2015. Full genome sequence of guinea fowl coronavirus associated with fulminating disease. *Virus Genes* 50, 514–517. <https://doi.org/10.1007/s11262-015-1183-z>
- Ducatez, M.F., The European Union COST Action FA1207, 2016. Recommendations for a standardized avian coronavirus (AvCoV) nomenclature: outcome from discussions within the framework of the European Union COST Action FA1207: “towards control of avian coronaviruses: strategies for vaccination, diagnosis and surveillance.” *Avian Pathology* 45, 602–603. <https://doi.org/10.1080/03079457.2016.1211834>

- Eckerle, L.D., Lu, X., Sperry, S.M., Choi, L., Denison, M.R., 2007. High fidelity of murine hepatitis virus replication is decreased in nsp14 exoribonuclease mutants. *J. Virol.* 81, 12135–12144. <https://doi.org/10.1128/JVI.01296-07>
- Eldemery, F., Joiner, K.S., Toro, H., van Santen, V.L., 2017. Protection against infectious bronchitis virus by spike ectodomain subunit vaccine. *Vaccine* 35, 5864–5871. <https://doi.org/10.1016/j.vaccine.2017.09.013>
- Fabricant, J., 1998. The early history of infectious bronchitis. *Avian Dis.* 42, 648–650.
- Fabricant, J., 1951. Studies on the diagnosis of Newcastle disease and infectious bronchitis. IV. The use of the serum neutralization test in the diagnosis of infectious bronchitis. *Cornell Vet* 41, 68–80.
- Felippe, P. a. N., da Silva, L.H.A., Santos, M.M. a. B., Spilki, F.R., Arns, C.W., 2010. Genetic diversity of avian infectious bronchitis virus isolated from domestic chicken flocks and coronaviruses from feral pigeons in Brazil between 2003 and 2009. *Avian Dis.* 54, 1191–1196. <https://doi.org/10.1637/9371-041510-Reg.1>
- Fernando, F.S., Kasmanas, T.C., Lopes, P.D., Montassier, H.J., Assayag, M.S., Jr., 2017. An infectious bronchitis virus Mass-like strain from Brazil causing kidney damage and tropism in a broiler flock. *Veterinary Record Case Reports* 5. <https://doi.org/10.1136/vetreccr-2017-000492>
- Ferreira, H.L., Pilz, D., Mesquita, L.G., Cardoso, T., 2003. Infectious bronchitis virus replication in the chicken embryo related cell line. *Avian Pathol.* 32, 413–417. <https://doi.org/10.1080/0307945031000121167>
- Fields, B.N., Knipe, D.M., Howley, P.M. (Eds.), 2013. *Fields virology*, 6th ed. ed. Wolters Kluwer Health/Lippincott Williams & Wilkins, Philadelphia.
- Gallardo, R.A., Santen, V.L. van, Toro, H., 2010. Host Intraspatial Selection of Infectious Bronchitis Virus Populations. *avdi* 54, 807–813. <https://doi.org/10.1637/9054-090809-Reg.1>
- Gallardo, R.A., van Santen, V.L., Toro, H., 2012. Effects of chicken anaemia virus and infectious bursal disease virus-induced immunodeficiency on infectious bronchitis virus replication and genotypic drift. *Avian Pathol.* 41, 451–458. <https://doi.org/10.1080/03079457.2012.702889>
- Geller, R., Estada, Ú., Peris, J.B., Andreu, I., Bou, J.-V., Garijo, R., Cuevas, J.M., Sabariego, R., Mas, A., Sanjuán, R., 2016. Highly heterogeneous mutation rates in the hepatitis C virus genome. *Nat Microbiol* 1, 16045. <https://doi.org/10.1038/nmicrobiol.2016.45>
- Gomaa, M.H., Barta, J.R., Ojkic, D., Yoo, D., 2008. Complete genomic sequence of turkey coronavirus. *Virus Res.* 135, 237–246. <https://doi.org/10.1016/j.virusres.2008.03.020>
- Gomaa, M.H., Yoo, D., Ojkic, D., Barta, J.R., 2009a. Infection with a pathogenic turkey coronavirus isolate negatively affects growth performance and intestinal morphology of young turkey poults in Canada. *Avian Pathol.* 38, 279–286. <https://doi.org/10.1080/03079450903055389>
- Gomaa, M.H., Yoo, D., Ojkic, D., Barta, J.R., 2009b. Virus shedding and serum antibody responses during experimental turkey coronavirus infections in young turkey poults. *Avian Pathol.* 38, 181–186. <https://doi.org/10.1080/03079450902751863>
- Gomes, D.E., Hirata, K.Y., Saheki, K., Rosa, A.C.G., Luvizotto, M.C.R., Cardoso, T.C., 2010. Pathology and tissue distribution of turkey coronavirus in experimentally infected chicks and turkey poults. *J. Comp. Pathol.* 143, 8–13. <https://doi.org/10.1016/j.jcpa.2009.12.012>
- Goodacre, N., Aljanahi, A., Nandakumar, S., Mikailov, M., Khan, A.S., 2018. A Reference Viral Database (RVDB) To Enhance Bioinformatics Analysis of High-Throughput Sequencing for Novel Virus Detection. *mSphere* 3. <https://doi.org/10.1128/mSphereDirect.00069-18>

- Gough, R.E., Cox, W.J., Winkler, C.E., Sharp, M.W., Spackman, D., 1996. Isolation and identification of infectious bronchitis virus from pheasants. *Vet. Rec.* 138, 208–209. <https://doi.org/10.1136/vr.138.9.208>
- Gough, R.E., Randall, C.J., Dagless, M., Alexander, D.J., Cox, W.J., Pearson, D., 1992. A “new” strain of infectious bronchitis virus infecting domestic fowl in Great Britain. *Vet. Rec.* 130, 493–494. <https://doi.org/10.1136/vr.130.22.493>
- Graham, R.L., Sims, A.C., Baric, R.S., Denison, M.R., 2006. The NSP2 Proteins of Mouse Hepatitis Virus and Sars Coronavirus are Dispensable for Viral Replication. *The Nidoviruses* 581, 67–72. https://doi.org/10.1007/978-0-387-33012-9_10
- Guionie, O., Courtillon, C., Allee, C., Maurel, S., Queguiner, M., Etteradossi, N., 2013. An experimental study of the survival of turkey coronavirus at room temperature and +4°C. *Avian Pathol.* 42, 248–252. <https://doi.org/10.1080/03079457.2013.779364>
- Guo, C.-T., Takahashi, N., Yagi, H., Kato, K., Takahashi, T., Yi, S.-Q., Chen, Y., Ito, T., Otsuki, K., Kida, H., Kawaoka, Y., Hidari, K.I.-P.J., Miyamoto, D., Suzuki, T., Suzuki, Y., 2007. The quail and chicken intestine have sialyl-galactose sugar chains responsible for the binding of influenza A viruses to human type receptors. *Glycobiology* 17, 713–724. <https://doi.org/10.1093/glycob/cwm038>
- Guy, J.S., 2017. Viral Enteric Infections, in: *Diseases of Poultry*, 13th Edition. John Wiley & Sons, Ltd, pp. 375–415. <https://doi.org/10.1002/9781119421481.ch12>
- Guy, J.S., Barnes, H.J., Smith, L.G., Breslin, J., 1997. Antigenic Characterization of a Turkey Coronavirus Identified in Poult Enteritis- and Mortality Syndrome-Affected Turkeys. *Avian Diseases* 41, 583–590. <https://doi.org/10.2307/1592148>
- Guy, J.S., Breslin, J.J., Breuhaus, B., Vivrette, S., Smith, L.G., 2000a. Characterization of a Coronavirus Isolated from a Diarrheic Foal. *Journal of Clinical Microbiology* 38, 4523–4526. <https://doi.org/10.1128/JCM.38.12.4523-4526.2000>
- Guy, J.S., Smith, L.G., Breslin, J.J., Vaillancourt, J.P., Barnes, H.J., 2000b. High Mortality and Growth Depression Experimentally Produced in Young Turkeys by Dual Infection with Enteropathogenic Escherichia coli and Turkey Coronavirus. *Avian Diseases* 44, 105–113. <https://doi.org/10.2307/1592513>
- Hackbart, M., Deng, X., Baker, S.C., 2020. Coronavirus endoribonuclease targets viral polyuridine sequences to evade activating host sensors. *Proc Natl Acad Sci U S A* 117, 8094–8103. <https://doi.org/10.1073/pnas.1921485117>
- Hewson, K.A., Ignjatovic, J., Browning, G.F., Devlin, J.M., Noormohammadi, A.H., 2011. Infectious bronchitis viruses with naturally occurring genomic rearrangement and gene deletion. *Arch. Virol.* 156, 245–252. <https://doi.org/10.1007/s00705-010-0850-6>
- Hofstad, M.S., Yoder, H.W., 1966. Avian infectious bronchitis--virus distribution in tissues of chicks. *Avian Dis.* 10, 230–239.
- Hopkins, S.R., Yoder, H.W., 1986. Reversion to virulence of chicken-passaged infectious bronchitis vaccine virus. *Avian Dis.* 30, 221–223.
- Hu, H., Jung, K., Vlasova, A.N., Chepngeno, J., Lu, Z., Wang, Q., Saif, L.J., 2015. Isolation and Characterization of Porcine Deltacoronavirus from Pigs with Diarrhea in the United States. *Journal of Clinical Microbiology* 53, 1537–1548. <https://doi.org/10.1128/JCM.00031-15>
- Hulo, C., de Castro, E., Masson, P., Bougueleret, L., Bairoch, A., Xenarios, I., Le Mercier, P., 2011. ViralZone: a knowledge resource to understand virus diversity. *Nucleic Acids Res* 39, D576–D582. <https://doi.org/10.1093/nar/gkq901>
- Hulswit, R.J.G., de Haan, C.A.M., Bosch, B.-J., 2016. Coronavirus Spike Protein and Tropism Changes. *Adv Virus Res* 96, 29–57. <https://doi.org/10.1016/bs.aivir.2016.08.004>
- Hyde, J.L., Diamond, M.S., 2015. Innate immune restriction and antagonism of viral RNA lacking 2'-O methylation. *Virology* 479, 66–74. <https://doi.org/10.1016/j.virol.2015.01.019>

- Imbert, I., Guillemot, J.-C., Bourhis, J.-M., Bussetta, C., Coutard, B., Egloff, M.-P., Ferron, F., Gorbalenya, A.E., Canard, B., 2006. A second, non-canonical RNA-dependent RNA polymerase in SARS Coronavirus. *The EMBO Journal* 25, 4933–4942. <https://doi.org/10.1038/sj.emboj.7601368>
- Isakov, O., Bordería, A.V., Golan, D., Hamenahem, A., Celniker, G., Yoffe, L., Blanc, H., Vignuzzi, M., Shomron, N., 2015. Deep sequencing analysis of viral infection and evolution allows rapid and detailed characterization of viral mutant spectrum. *Bioinformatics* 31, 2141–2150. <https://doi.org/10.1093/bioinformatics/btv101>
- Ismail, M.M., Tang, A.Y., Saif, Y.M., 2003. Pathogenicity of turkey coronavirus in turkeys and chickens. *Avian Dis.* 47, 515–522. <https://doi.org/10.1637/5917>
- Ivanov, K.A., Ziebuhr, J., 2004. Human Coronavirus 229E Nonstructural Protein 13: Characterization of Duplex-Unwinding, Nucleoside Triphosphatase, and RNA 5'-Triphosphatase Activities. *J Virol* 78, 7833–7838. <https://doi.org/10.1128/JVI.78.14.7833-7838.2004>
- Jabara, C.B., Jones, C.D., Roach, J., Anderson, J.A., Swanstrom, R., 2011. Accurate sampling and deep sequencing of the HIV-1 protease gene using a Primer ID. *PNAS* 108, 20166–20171. <https://doi.org/10.1073/pnas.1110064108>
- Jackwood, M.W., 2012. Review of infectious bronchitis virus around the world. *Avian Dis.* 56, 634–641. <https://doi.org/10.1637/10227-043012-Review.1>
- Jackwood, M.W., De Wit, S.J.J., 2013. Infectious Bronchitis, in: *Diseases of Poultry*. John Wiley & Sons, Ltd, pp. 139–159. <https://doi.org/10.1002/9781119421481.ch4>
- Jayaram, J., Youn, S., Collisson, E.W., 2005. The virion N protein of infectious bronchitis virus is more phosphorylated than the N protein from infected cell lysates. *Virology* 339, 127–135. <https://doi.org/10.1016/j.virol.2005.04.029>
- Jin, X., Chen, Y., Sun, Y., Zeng, C., Wang, Y., Tao, J., Wu, A., Yu, X., Zhang, Z., Tian, J., Guo, D., 2013. Characterization of the guanine-N7 methyltransferase activity of coronavirus nsp14 on nucleotide GTP. *Virus Research* 176, 45–52. <https://doi.org/10.1016/j.virusres.2013.05.001>
- Jindal, N., Mor, S.K., Goyal, S.M., 2014. Enteric viruses in turkey enteritis. *Virusdisease* 25, 173–185. <https://doi.org/10.1007/s13337-014-0198-8>
- Jindal, N., Patnayak, D.P., Chander, Y., Ziegler, A.F., Goyal, S.M., 2010. Detection and molecular characterization of enteric viruses from poult enteritis syndrome in turkeys. *Poult. Sci.* 89, 217–226. <https://doi.org/10.3382/ps.2009-00424>
- Jindal, N., Patnayak, D.P., Ziegler, A.F., Lago, A., Goyal, S.M., 2009. Experimental reproduction of poult enteritis syndrome: Clinical findings, growth response, and microbiology. *Poult Sci* 88, 949–958. <https://doi.org/10.3382/ps.2008-00490>
- Johnson, M.A., Pooley, C., Ignjatovic, J., Tyack, S.G., 2003. A recombinant fowl adenovirus expressing the S1 gene of infectious bronchitis virus protects against challenge with infectious bronchitis virus. *Vaccine* 21, 2730–2736. [https://doi.org/10.1016/s0264-410x\(03\)00227-5](https://doi.org/10.1016/s0264-410x(03)00227-5)
- Jonsson, L.G., Engstrom, B.E., 1986. Immunohistochemical detection of infectious bursal disease and infectious bronchitis viral antigens in fixed, paraffin-embedded chicken tissues. *Avian Pathol.* 15, 385–393. <https://doi.org/10.1080/03079458608436301>
- Jordan, B., 2017. Vaccination against infectious bronchitis virus: A continuous challenge. *Vet. Microbiol.* 206, 137–143. <https://doi.org/10.1016/j.vetmic.2017.01.002>
- Keep, S., Bickerton, E., Britton, P. (Eds.), 2015. Chapitre 11: Partial Purification of IBV and Subsequent Isolation of Viral RNA for Next-Generation Sequencing, in: *Coronaviruses: Methods and Protocols*, Methods in Molecular Biology. Humana Press.

- Keep, S.M., Bickerton, E., Britton, P., 2015. Partial purification of IBV and subsequent isolation of viral RNA for next-generation sequencing. *Methods Mol. Biol.* 1282, 109–112. https://doi.org/10.1007/978-1-4939-2438-7_11
- Khanh, N.P., Tan, S.W., Yeap, S.K., Lee, H.J., Choi, K.S., Hair-Bejo, M., Bich, T.N., Omar, A.R., 2018. Comparative Pathogenicity of Malaysian QX-like and Variant Infectious Bronchitis Virus Strains in Chickens at Different Age of Exposure to the Viruses. *Journal of Comparative Pathology* 161, 43–54. <https://doi.org/10.1016/j.jcpa.2018.04.006>
- King, A.M., Lefkowitz, E., Adams, M.J., Carstens, E.B., 2011. *Virus Taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses*. Elsevier.
- Kint, J., Fernandez-Gutierrez, M., Maier, H.J., Britton, P., Langereis, M.A., Koumans, J., Wiegertjes, G.F., Forlenza, M., 2015. Activation of the chicken type I interferon response by infectious bronchitis coronavirus. *J. Virol.* 89, 1156–1167. <https://doi.org/10.1128/JVI.02671-14>
- Koboldt, D.C., Zhang, Q., Larson, D.E., Shen, D., McLellan, M.D., Lin, L., Miller, C.A., Mardis, E.R., Ding, L., Wilson, R.K., 2012. VarScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* 22, 568–576. <https://doi.org/10.1101/gr.129684.111>
- Kuchipudi, S.V., Nelli, R., White, G.A., Bain, M., Chang, K.C., Dunham, S., 2009. Differences in influenza virus receptors in chickens and ducks: Implications for interspecies transmission. *J Mol Genet Med* 3, 143–151.
- Kusters, J.G., Jager, E.J., Lenstra, J.A., Koch, G., Posthumus, W.P., Meloen, R.H., Zeijst, B.A. van der, 1989. Analysis of an immunodominant region of infectious bronchitis virus. *The Journal of Immunology* 143, 2692–2698.
- Laconi, A., van Beurden, S.J., Berends, A.J., Krämer-Kühl, A., Jansen, C.A., Spekrijse, D., Chénard, G., Philipp, H.-C., Mundt, E., Rottier, P.J.M., Hélène Verheije, M., 2018. Deletion of accessory genes 3a, 3b, 5a or 5b from avian coronavirus infectious bronchitis virus induces an attenuated phenotype both in vitro and in vivo. *J. Gen. Virol.* <https://doi.org/10.1099/jgv.0.001130>
- Lai, Z., Markovets, A., Ahdesmaki, M., Chapman, B., Hofmann, O., McEwen, R., Johnson, J., Dougherty, B., Barrett, J.C., Dry, J.R., 2016. VarDict: a novel and versatile variant caller for next-generation sequencing in cancer research. *Nucleic Acids Res* 44, e108–e108. <https://doi.org/10.1093/nar/gkw227>
- Lambrechts, C., Pensaert, M., Ducatelle, R., 1993. Challenge experiments to evaluate cross-protection induced at the trachea and kidney level by vaccine strains and Belgian nephropathogenic isolates of avian infectious bronchitis virus. *Avian Pathol.* 22, 577–590. <https://doi.org/10.1080/03079459308418945>
- Lee, C.-W., Brown, C., Jackwood, M.W., 2002. Tissue distribution of avian infectious bronchitis virus following in ovo inoculation of chicken embryos examined by in situ hybridization with antisense digoxigenin-labeled universal riboprobe. *J. Vet. Diagn. Invest.* 14, 377–381. <https://doi.org/10.1177/104063870201400503>
- Lee, C.W., Hilt, D.A., Jackwood, M.W., 2001. Identification and analysis of the Georgia 98 serotype, a new serotype of infectious bronchitis virus. *Avian Dis.* 45, 164–172.
- Legnardi, M., Franzo, G., Koutoulis, K.C., Wiśniewski, M., Catelli, E., Tucciarone, C.M., Cecchinato, M., 2019. Vaccine or field strains: the jigsaw pattern of infectious bronchitis virus molecular epidemiology in Poland. *Poult. Sci.* <https://doi.org/10.3382/ps/pez473>
- Lei, J., Kusov, Y., Hilgenfeld, R., 2018. Nsp3 of coronaviruses: Structures and functions of a large multi-domain protein. *Antiviral Res* 149, 58–74. <https://doi.org/10.1016/j.antiviral.2017.11.001>

- Li, C., Chng, K.R., Boey, E.J.H., Ng, A.H.Q., Wilm, A., Nagarajan, N., 2016. INC-Seq: accurate single molecule reads using nanopore sequencing. *Gigascience* 5. <https://doi.org/10.1186/s13742-016-0140-7>
- Li, F., 2015. Receptor Recognition Mechanisms of Coronaviruses: a Decade of Structural Studies. *Journal of Virology* 89, 1954–1964. <https://doi.org/10.1128/JVI.02615-14>
- Li, L., Fu, F., Guo, S., Wang, H., He, X., Xue, M., Yin, L., Feng, L., Liu, P., 2018. Porcine Intestinal Enteroids: a New Model for Studying Enteric Coronavirus Porcine Epidemic Diarrhea Virus Infection and the Host Innate Response. *Journal of Virology* 93. <https://doi.org/10.1128/JVI.01682-18>
- Lin, T.L., Loa, C.C., Tsai, S.C., Wu, C.C., Bryan, T.A., Leon Thacker, H., Hooper, T., Schrader, D., 2002. Characterization of turkey coronavirus from turkey poult with acute enteritis. *Vet Microbiol* 84, 179–186. [https://doi.org/10.1016/S0378-1135\(01\)00447-3](https://doi.org/10.1016/S0378-1135(01)00447-3)
- Linares-Fernández, S., Lacroix, C., Exposito, J.-Y., Verrier, B., 2020. Tailoring mRNA Vaccine to Balance Innate/Adaptive Immune Response. *Trends in Molecular Medicine* 26, 311–323. <https://doi.org/10.1016/j.molmed.2019.10.002>
- Liu, D.X., Cavanagh, D., Green, P., Inglis, S.C., 1991. A polycistronic mRNA specified by the coronavirus infectious bronchitis virus. *Virology* 184, 531–544. [https://doi.org/10.1016/0042-6822\(91\)90423-9](https://doi.org/10.1016/0042-6822(91)90423-9)
- Liu, D. X., Inglis, S.C., 1992. Internal entry of ribosomes on a tricistronic mRNA encoded by infectious bronchitis virus. *J. Virol.* 66, 6143–6154.
- Liu, D.X., Inglis, S.C., 1992. Identification of two new polypeptides encoded by mRNA5 of the coronavirus infectious bronchitis virus. *Virology* 186, 342–347. [https://doi.org/10.1016/0042-6822\(92\)90094-6](https://doi.org/10.1016/0042-6822(92)90094-6)
- Liu, G., Lv, L., Yin, L., Li, X., Luo, D., Liu, K., Xue, C., Cao, Y., 2013. Assembly and immunogenicity of coronavirus-like particles carrying infectious bronchitis virus M and S proteins. *Vaccine* 31, 5524–5530. <https://doi.org/10.1016/j.vaccine.2013.09.024>
- Liu, S., Kong, X., 2004. A new genotype of nephropathogenic infectious bronchitis virus circulating in vaccinated and non-vaccinated flocks in China. *Avian Pathol.* 33, 321–327. <https://doi.org/10.1080/0307945042000220697>
- Lou, D.I., Hussmann, J.A., McBee, R.M., Acevedo, A., Andino, R., Press, W.H., Sawyer, S.L., 2013. High-throughput DNA sequencing errors are reduced by orders of magnitude using circle sequencing. *PNAS* 110, 19872–19877. <https://doi.org/10.1073/pnas.1319590110>
- Lu, I.-N., Muller, C.P., He, F.Q., 2020. Applying next-generation sequencing to unravel the mutational landscape in viral quasispecies. *Virus Res* 283, 197963. <https://doi.org/10.1016/j.virusres.2020.197963>
- Luginbuhl, R.E., 1952. Infectious Bronchitis Control. *World's Poultry Science Journal* 8, 182–183. <https://doi.org/10.1079/WPS19520031>
- Lukert, P.D., 1966. Immunofluorescence of avian infectious bronchitis virus in primary chicken embryo kidney, liver, lung, and fibroblast cell cultures. *Arch Gesamte Virusforsch* 19, 265–272. <https://doi.org/10.1007/bf01241849>
- MacNaughton, M.R., 1978. The genomes of three coronaviruses. *FEBS Lett.* 94, 191–194. [https://doi.org/10.1016/0014-5793\(78\)80935-1](https://doi.org/10.1016/0014-5793(78)80935-1)
- Macnaughton, M.R., Davies, H.A., 1980. Two Particle Types of Avian Infectious Bronchitis Virus. *Journal of General Virology* 47, 365–372. <https://doi.org/10.1099/0022-1317-47-2-365>
- Macneughton, M.R., Davies, H.A., 1978. Ribonucleoprotein-like structures from coronavirus particles. *J. Gen. Virol.* 39, 545–549. <https://doi.org/10.1099/0022-1317-39-3-545>
- Madu, I.G., Chu, V.C., Lee, H., Regan, A.D., Bauman, B.E., Whittaker, G.R., 2007. Heparan Sulfate Is a Selective Attachment Factor for the Avian Coronavirus Infectious Bronchitis

- Virus Beaudette. *avdi* 51, 45–51. [https://doi.org/10.1637/0005-2086\(2007\)051\[0045:HSIASA\]2.0.CO;2](https://doi.org/10.1637/0005-2086(2007)051[0045:HSIASA]2.0.CO;2)
- Maier, H.J., Hawes, P.C., Keep, S.M., Britton, P., 2014. Spherules and IBV. *Bioengineered* 5, 288–292. <https://doi.org/10.4161/bioe.29323>
- Martini, M.C., Caserta, L.C., Dos Santos, M.M.A.B., Barnabé, A.C.S., Durães-Carvalho, R., Padilla, M.A., Simão, R.M., Rizotto, L.S., Simas, P.V.M., Bastos, J.C.S., Cardoso, T.C., Felipe, P.A.N., Ferreira, H.L., Arns, C.W., 2018. Avian coronavirus isolated from a pigeon sample induced clinical disease, tracheal ciliostasis, and a high humoral response in day-old chicks. *Avian Pathol.* 47, 286–293. <https://doi.org/10.1080/03079457.2018.1442557>
- Mase, M., Inoue, T., Yamaguchi, S., Imada, T., 2008. Existence of avian infectious bronchitis virus with a European-prevalent 4/91 genotype in Japan. *J. Vet. Med. Sci.* 70, 1341–1344. <https://doi.org/10.1292/jvms.70.1341>
- Masters, P.S., 2019. Coronavirus genomic RNA packaging. *Virology* 537, 198–207. <https://doi.org/10.1016/j.virol.2019.08.031>
- Maurel, S., Toquin, D., Briand, F.X., Queguiner, M., Allee, C., Bertin, J., Ravillion, L., Retaux, C., Turblin, V., Morvan, H., Eterradosi, N., 2011. First full-length sequences of the S gene of European isolates reveal further diversity among turkey coronaviruses. *Avian Pathol.* 40, 179–189. <https://doi.org/10.1080/03079457.2011.551936>
- McBride, R., van Zyl, M., Fielding, B.C., 2014. The Coronavirus Nucleocapsid Is a Multifunctional Protein. *Viruses* 6, 2991–3018. <https://doi.org/10.3390/v6082991>
- McIntosh, K., 1974. Coronaviruses: A Comparative Review, in: Arber, W., Haas, R., Henle, W., Hofschneider, P.H., Jerne, N.K., Koldovský, P., Koprowski, H., Maaløe, O., Rott, R., Schweiger, H.G., Sela, M., Syruček, L., Vogt, P.K., Wecker, E. (Eds.), *Current Topics in Microbiology and Immunology / Ergebnisse Der Mikrobiologie Und Immunitätsforschung*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 85–129. https://doi.org/10.1007/978-3-642-65775-7_3
- McKinley, E.T., Hilt, D.A., Jackwood, M.W., 2008. Avian coronavirus infectious bronchitis attenuated live vaccines undergo selection of subpopulations and mutations following vaccination. *Vaccine* 26, 1274–1284. <https://doi.org/10.1016/j.vaccine.2008.01.006>
- Miguel, B., Pharr, G.T., Wang, C., 2002. The role of feline aminopeptidase N as a receptor for infectious bronchitis virus. Brief review. *Arch. Virol.* 147, 2047–2056. <https://doi.org/10.1007/s00705-002-0888-1>
- Mihindukulasuriya, K.A., Wu, G., Leger, J.S., Nordhausen, R.W., Wang, D., 2008. Identification of a Novel Coronavirus from a Beluga Whale by Using a Panviral Microarray. *Journal of Virology* 82, 5084–5088. <https://doi.org/10.1128/JVI.02722-07>
- Molenaar, R.J., Dijkman, R., De Wit, S.J.J., 2020. Characterization of infectious bronchitis virus D181, a new serotype (GII-2). *Avian Pathol.* 1–8. <https://doi.org/10.1080/03079457.2020.1713987>
- Mork, A.-K., Hesse, M., Abd El Rahman, S., Rautenschlein, S., Herrler, G., Winter, C., 2014. Differences in the tissue tropism to chicken oviduct epithelial cells between avian coronavirus IBV strains QX and B1648 are not related to the sialic acid binding properties of their spike proteins. *Vet. Res.* 45, 67. <https://doi.org/10.1186/1297-9716-45-67>
- Moura-Alvarez, J., Chacon, J.V., Scanavini, L.S., Nuñez, L.F.N., Astolfi-Ferreira, C.S., Jones, R.C., Piantino Ferreira, A.J., 2013. Enteric viruses in Brazilian turkey flocks: single and multiple virus infection frequency according to age and clinical signs of intestinal disease. *Poult. Sci.* 92, 945–955. <https://doi.org/10.3382/ps.2012-02849>
- Moura-Alvarez, J., Nuñez, L.F.N., Astolfi-Ferreira, C.S., Knöbl, T., Chacón, J.L., Moreno, A.M., Jones, R.C., Ferreira, A.J.P., 2014. Detection of enteric pathogens in Turkey flocks affected

- with severe enteritis, in Brazil. *Trop Anim Health Prod* 46, 1051–1058. <https://doi.org/10.1007/s11250-014-0612-7>
- Nakagawa, K., Lokugamage, K.G., Makino, S., 2016. Viral and Cellular mRNA Translation in Coronavirus-Infected Cells. *Adv Virus Res* 96, 165–192. <https://doi.org/10.1016/bs.aivir.2016.08.001>
- Narayanan, K., Ramirez, S.I., Lokugamage, K.G., Makino, S., 2015. Coronavirus nonstructural protein 1: Common and distinct functions in the regulation of host and viral gene expression. *Virus Res* 202, 89–100. <https://doi.org/10.1016/j.virusres.2014.11.019>
- Neuman, B.W., Kiss, G., Kunding, A.H., Bhella, D., Baksh, M.F., Connelly, S., Droese, B., Klaus, J.P., Makino, S., Sawicki, S.G., Siddell, S.G., Stamou, D.G., Wilson, I.A., Kuhn, P., Buchmeier, M.J., 2011. A structural analysis of M protein in coronavirus assembly and morphology. *J Struct Biol* 174, 11–22. <https://doi.org/10.1016/j.jsb.2010.11.021>
- Ng, L.F., Liu, D.X., 1998. Identification of a 24-kDa polypeptide processed from the coronavirus infectious bronchitis virus 1a polyprotein by the 3C-like proteinase and determination of its cleavage sites. *Virology* 243, 388–395. <https://doi.org/10.1006/viro.1998.9058>
- Nix, W.A., Troeber, D.S., Kingham, B.F., Keeler, C.L., Gelb, J., 2000. Emergence of subtype strains of the Arkansas serotype of infectious bronchitis virus in Delmarva broiler chickens. *Avian Dis.* 44, 568–581.
- Oade, M.S., Keep, S., Freimanis, G.L., Orton, R.J., Britton, P., Hammond, J.A., Bickerton, E., 2019. Attenuation of Infectious Bronchitis Virus in Eggs Results in Different Patterns of Genomic Variation across Multiple Replicates. *J. Virol.* 93. <https://doi.org/10.1128/JVI.00492-19>
- Pakpinyo, S., Ley, D.H., Barnes, H.J., Vaillancourt, J.P., Guy, J.S., 2003. Enhancement of enteropathogenic *Escherichia coli* pathogenicity in young turkeys by concurrent turkey coronavirus infection. *Avian Dis.* 47, 396–405. [https://doi.org/10.1637/0005-2086\(2003\)047\[0396:EOEECP\]2.0.CO;2](https://doi.org/10.1637/0005-2086(2003)047[0396:EOEECP]2.0.CO;2)
- Panigrahy, B., Naqi, S.A., Hall, C.F., 1973. Isolation and characterization of viruses associated with transmissible enteritis (bluecomb) of turkeys. *Avian Dis.* 17, 430–438.
- Pensaert, M., Lambrechts, C., 1994. Vaccination of chickens against a Belgian nephropathogenic strain of infectious bronchitis virus B1648 using attenuated homologous and heterologous strains. *Avian Pathol.* 23, 631–641. <https://doi.org/10.1080/03079459408419033>
- Pereira, N.A., Alessi, A.C., Montassier, H.J., Pereira, R.J.G., Taniwaki, S.A., Botosso, V.F., Rui, B.R., Richtzenhain, L.J., 2018. Gonadal pathogenicity of an infectious bronchitis virus strain from the Massachusetts genotype. *Braz J Microbiol* 50, 313–320. <https://doi.org/10.1007/s42770-018-0007-4>
- Pickett, B.E., Greer, D.S., Zhang, Y., Stewart, L., Zhou, L., Sun, G., Gu, Z., Kumar, S., Zaremba, S., Larsen, C.N., Jen, W., Klem, E.B., Scheuermann, R.H., 2012. Virus Pathogen Database and Analysis Resource (ViPR): A Comprehensive Bioinformatics Database and Analysis Resource for the Coronavirus Research Community. *Viruses* 4, 3209–3226. <https://doi.org/10.3390/v4113209>
- Promkuntod, N., van Eijndhoven, R.E.W., de Vrieze, G., Gröne, A., Verheije, M.H., 2014. Mapping of the receptor-binding domain and amino acids critical for attachment in the spike protein of avian coronavirus infectious bronchitis virus. *Virology* 448, 26–32. <https://doi.org/10.1016/j.virol.2013.09.018>
- Promkuntod, N., Wickramasinghe, I.N.A., de Vrieze, G., Gröne, A., Verheije, M.H., 2013. Contributions of the S2 spike ectodomain to attachment and host range of infectious bronchitis virus. *Virus Research* 177, 127–137. <https://doi.org/10.1016/j.virusres.2013.09.006>

- Qian, D.H., Zhu, G.J., Wu, L.Z., Hua, G.X., 2006. Isolation and characterization of a coronavirus from pigeons with pancreatitis. *Am. J. Vet. Res.* 67, 1575–1579. <https://doi.org/10.2460/ajvr.67.9.1575>
- Quail, M.A., Smith, M., Coupland, P., Otto, T.D., Harris, S.R., Connor, T.R., Bertoni, A., Swerdlow, H.P., Gu, Y., 2012. A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics* 13, 341. <https://doi.org/10.1186/1471-2164-13-341>
- Rahman, S.A.E., El-Kenawy, A.A., Neumann, U., Herrler, G., Winter, C., 2009. Comparative analysis of the sialic acid binding activity and the tropism for the respiratory epithelium of four different strains of avian infectious bronchitis virus. *Avian Pathology* 38, 41–45. <https://doi.org/10.1080/03079450802632049>
- Ritchie, A.E., Deshmukh, D.R., Larsen, C.T., Pomeroy, B.S., 1973. Electron microscopy of coronavirus-like particles characteristic of turkey bluecomb disease. *Avian Dis.* 17, 546–558.
- Ruch, T.R., Machamer, C.E., 2011. The hydrophobic domain of infectious bronchitis virus E protein alters the host secretory pathway and is important for release of infectious virus. *J. Virol.* 85, 675–685. <https://doi.org/10.1128/JVI.01570-10>
- Sadri, N., Ghalyanchilangeroudi, A., Fallah Mehrabadi, M.H., Hosseini, H., Shayeganmehr, A., Sediqian, M.S., Jabbarifakhr, M., Hamdan, A.M., Mousavi, F.S., 2019. Genotyping of avian infectious bronchitis virus in Afghanistan (2016-2017): the first report. *Iran J Vet Res* 20, 60–63.
- Saiada, F., Gallardo, R.A., Shivaprasad, H.L., Corsiglia, C., Van Santen, V.L., 2020. Intestinal Tropism of an Infectious Bronchitis Virus Isolate Not Explained by Spike Protein Binding Specificity. *Avian Dis.* 64, 23–35. <https://doi.org/10.1637/0005-2086-64.1.23>
- Sandmann, S., Graaf, A.O. de, Karimi, M., Reijden, B.A. van der, Hellström-Lindberg, E., Jansen, J.H., Dugas, M., 2017. Evaluating Variant Calling Tools for Non-Matched Next-Generation Sequencing Data. *Scientific Reports* 7, 43169. <https://doi.org/10.1038/srep43169>
- Sawicki, S.G., Sawicki, D.L., 1995. Coronaviruses use Discontinuous Extension for Synthesis of Subgenome-Length Negative Strands, in: Talbot, P.J., Levy, G.A. (Eds.), *Corona- and Related Viruses: Current Concepts in Molecular Biology and Pathogenesis*, *Advances in Experimental Medicine and Biology*. Springer US, Boston, MA, pp. 499–506. https://doi.org/10.1007/978-1-4615-1899-0_79
- Schalk, Hawn, 1931. First Report of Infectious Bronchitis in North Dakota - An Apparently New Respiratory Disease of Chicks. *Journal of the American Veterinary Medical Association* 78, 413–422.
- Schochetman, G., Stevens, R.H., Simpson, R.W., 1977. Presence of infectious polyadenylated RNA in coronavirus avian bronchitis virus. *Virology* 77, 772–782. [https://doi.org/10.1016/0042-6822\(77\)90498-6](https://doi.org/10.1016/0042-6822(77)90498-6)
- Sethna, P.B., Hung, S.L., Brian, D.A., 1989. Coronavirus subgenomic minus-strand RNAs and the potential for mRNA replicons. *Proceedings of the National Academy of Sciences* 86, 5626–5630. <https://doi.org/10.1073/pnas.86.14.5626>
- Shan, D., Fang, S., Han, Z., Ai, H., Zhao, W., Chen, Y., Jiang, L., Liu, S., 2018. Effects of hypervariable regions in spike protein on pathogenicity, tropism, and serotypes of infectious bronchitis virus. *Virus Research* 250, 104–113. <https://doi.org/10.1016/j.virusres.2018.04.013>
- Shang, J., Zheng, Y., Yang, Y., Liu, C., Geng, Q., Luo, C., Zhang, W., Li, F., 2018. Cryo-EM structure of infectious bronchitis coronavirus spike protein reveals structural and functional evolution of coronavirus spike proteins. *PLoS Pathog.* 14, e1007009. <https://doi.org/10.1371/journal.ppat.1007009>

- Shen, S., Wen, Z.L., Liu, D.X., 2003. Emergence of a coronavirus infectious bronchitis virus mutant with a truncated 3b gene: functional characterization of the 3b protein in pathogenesis and replication. *Virology* 311, 16–27. [https://doi.org/10.1016/s0042-6822\(03\)00117-x](https://doi.org/10.1016/s0042-6822(03)00117-x)
- Sigrist, B., Tobler, K., Schybli, M., Konrad, L., Stöckli, R., Cattoli, G., Lüscho, D., Hafez, H.M., Britton, P., Hoop, R.K., Vöglin, A., 2012. Detection of Avian coronavirus infectious bronchitis virus type QX infection in Switzerland. *J. Vet. Diagn. Invest.* 24, 1180–1183. <https://doi.org/10.1177/1040638712463692>
- Snijder, E.J., Decroly, E., Ziebuhr, J., 2016. The Nonstructural Proteins Directing Coronavirus RNA Synthesis and Processing. *Adv. Virus Res.* 96, 59–126. <https://doi.org/10.1016/bs.aivir.2016.08.008>
- Sola, I., Almazán, F., Zúñiga, S., Enjuanes, L., 2015. Continuous and Discontinuous RNA Synthesis in Coronaviruses. *Annu Rev Virol* 2, 265–288. <https://doi.org/10.1146/annurev-virology-100114-055218>
- Spackman, D., Cameron, I.R., 1983. Isolation of infectious bronchitis virus from pheasants. *Veterinary Record* 113, 354–355. <https://doi.org/10.1136/vr.113.15.354>
- Sturman, L.S., Holmes, K.V., 1983. The Molecular Biology of Coronaviruses, in: Lauffer, M.A., Maramorosch, K. (Eds.), *Advances in Virus Research*. Academic Press, pp. 35–112. [https://doi.org/10.1016/S0065-3527\(08\)60721-6](https://doi.org/10.1016/S0065-3527(08)60721-6)
- Subissi, L., Posthuma, C.C., Collet, A., Zevenhoven-Dobbe, J.C., Gorbalenya, A.E., Decroly, E., Snijder, E.J., Canard, B., Imbert, I., 2014. One severe acute respiratory syndrome coronavirus protein complex integrates processive RNA polymerase and exonuclease activities. *PNAS* 111, E3900–E3909. <https://doi.org/10.1073/pnas.1323705111>
- Tan, Y.W., Fung, T.S., Shen, H., Huang, M., Liu, D.X., 2018. Coronavirus infectious bronchitis virus non-structural proteins 8 and 12 form stable complex independent of the non-translated regions of viral RNA and other viral proteins. *Virology* 513, 75–84. <https://doi.org/10.1016/j.virol.2017.10.004>
- te Velthuis, A.J.W., van den Worm, S.H.E., Snijder, E.J., 2012. The SARS-coronavirus nsp7+nsp8 complex is a unique multimeric RNA polymerase capable of both de novo initiation and primer extension. *Nucleic Acids Res* 40, 1737–1747. <https://doi.org/10.1093/nar/gkr893>
- Teixeira, M.C.B., Luvizotto, M.C.R., Ferrari, H.F., Mendes, A.R., da Silva, S.E.L., Cardoso, T.C., 2007. Detection of turkey coronavirus in commercial turkey poults in Brazil. *Avian Pathol.* 36, 29–33. <https://doi.org/10.1080/03079450601102939>
- Terregino, C., Toffan, A., Beato, M.S., De Nardi, R., Vascellari, M., Meini, A., Ortali, G., Mancin, M., Capua, I., 2008. Pathogenicity of a QX strain of infectious bronchitis virus in specific pathogen free and commercial broiler chickens, and evaluation of protection induced by a vaccination programme based on the Ma5 and 4/91 serotypes. *Avian Pathol.* 37, 487–493. <https://doi.org/10.1080/03079450802356938>
- The World Bank, 2011. *World livestock disease atlas : a quantitative analysis of global animal health data (2006-2009)* (No. 66859). The World Bank.
- Toro, H., Pennington, D., Gallardo, R.A., van Santen, V.L., van Ginkel, F.W., Zhang, J., Joiner, K.S., 2012. Infectious bronchitis virus subpopulations in vaccinated chickens after challenge. *Avian Dis.* 56, 501–508. <https://doi.org/10.1637/9982-110811-Reg.1>
- Torres, C.A., Hora, A.S., Tonietti, P.O., Taniwaki, S.A., Cecchinato, M., Villarreal, L.Y.B., Brandão, P.E., 2016. Gammacoronavirus and Deltacoronavirus in Quail. *Avian Dis.* 60, 656–661. <https://doi.org/10.1637/11412-032316-Reg.1>
- Torres, C.A., Listorti, V., Lupini, C., Franzo, G., Drigo, M., Catelli, E., Brandão, P.E., Cecchinato, M., 2017. Gamma and Deltacoronaviruses in quail and pheasants from Northern Italy. *Poult. Sci.* 96, 717–722. <https://doi.org/10.3382/ps/pew332>

- Torres, C.A., Villarreal, L.Y.B., Ayres, G.R.R., Richtzenhain, L.J., Brandão, P.E., 2013. An Avian coronavirus in quail with respiratory and reproductive signs. *Avian Dis.* 57, 295–299. <https://doi.org/10.1637/10412-100412-Reg.1>
- Tvarogová, J., Madhugiri, R., Bylapudi, G., Ferguson, L.J., Karl, N., Ziebuhr, J., 2019. Identification and Characterization of a Human Coronavirus 229E Nonstructural Protein 8-Associated RNA 3'-Terminal Adenylyltransferase Activity. *Journal of Virology* 93. <https://doi.org/10.1128/JVI.00291-19>
- Tyrrell, D.A., Almeida, J.D., Berry, D.M., Hamre, D., Hofstad, M.S., Mallucci, L., McIntosh, K., 1968. Virology - Coronaviruses. *Nature* 220, 650–650. <https://doi.org/10.1038/220650a0>
- Valastro, V., Holmes, E.C., Britton, P., Fusaro, A., Jackwood, M.W., Cattoli, G., Monne, I., 2016. S1 gene-based phylogeny of infectious bronchitis virus: An attempt to harmonize virus classification. *Infect. Genet. Evol.* 39, 349–364. <https://doi.org/10.1016/j.meegid.2016.02.015>
- van Beurden, S.J., Berends, A.J., Krämer-Kühl, A., Spekrijse, D., Chenard, G., Philipp, H.-C., Mundt, E., Rottier, P.J.M., Verheije, M.H., 2018. Recombinant live attenuated avian coronavirus vaccines with deletions in the accessory genes 3ab and/or 5ab protect against infectious bronchitis in chickens. *Vaccine* 36, 1085–1092. <https://doi.org/10.1016/j.vaccine.2018.01.017>
- van Santen, V.L., Toro, H., 2008. Rapid selection in chickens of subpopulations within ArkDPI-derived infectious bronchitis virus vaccines. *Avian Pathol.* 37, 293–306. <https://doi.org/10.1080/03079450802043783>
- Vignuzzi, M., Wendt, E., Andino, R., 2008. Engineering attenuated virus vaccines by controlling replication fidelity. *Nature Medicine* 14, 154. <https://doi.org/10.1038/nm1726>
- Wang, H., Yuan, X., Sun, Y., Mao, X., Meng, C., Tan, L., Song, C., Qiu, X., Ding, C., Liao, Y., 2019. Infectious bronchitis virus entry mainly depends on clathrin mediated endocytosis and requires classical endosomal/lysosomal system. *Virology* 528, 118–136. <https://doi.org/10.1016/j.virol.2018.12.012>
- Watkins, H., Reeve, P., Alexander, D.J., 1975. The ribonucleic acid of infectious bronchitis virus. *Arch. Virol.* 47, 279–286. <https://doi.org/10.1007/bf01317815>
- Westerbeck, J.W., Machamer, C.E., 2019. The Infectious Bronchitis Coronavirus Envelope Protein Alters Golgi pH To Protect the Spike Protein and Promote the Release of Infectious Virus. *J. Virol.* 93. <https://doi.org/10.1128/JVI.00015-19>
- Westerbeck, J.W., Machamer, C.E., 2015. A Coronavirus E Protein Is Present in Two Distinct Pools with Different Effects on Assembly and the Secretory Pathway. *J. Virol.* 89, 9313–9323. <https://doi.org/10.1128/JVI.01237-15>
- Wickramasinghe, I.N.A., de Vries, R.P., Weerts, E. a. W.S., van Beurden, S.J., Peng, W., McBride, R., Ducatez, M., Guy, J., Brown, P., Etteradossi, N., Gröne, A., Paulson, J.C., Verheije, M.H., 2015. Novel Receptor Specificity of Avian Gammacoronaviruses That Cause Enteritis. *J. Virol.* 89, 8783–8792. <https://doi.org/10.1128/JVI.00745-15>
- Wickramasinghe, I.N.A., van Beurden, S.J., Weerts, E. a. W.S., Verheije, M.H., 2014. The avian coronavirus spike protein. *Virus Res.* 194, 37–48. <https://doi.org/10.1016/j.virusres.2014.10.009>
- Wilm, A., Aw, P.P.K., Bertrand, D., Yeo, G.H.T., Ong, S.H., Wong, C.H., Khor, C.C., Petric, R., Hibberd, M.L., Nagarajan, N., 2012. LoFreq: a sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. *Nucleic Acids Res* 40, 11189–11201. <https://doi.org/10.1093/nar/gks918>
- Winter, C., Herrler, G., Neumann, U., 2008. Infection of the tracheal epithelium by infectious bronchitis virus is sialic acid dependent. *Microbes and Infection* 10, 367–373. <https://doi.org/10.1016/j.micinf.2007.12.009>

- Winter, C., Schwegmann-Wessels, C., Cavanagh, D., Neumann, U., Herrler, G., 2006. Sialic acid is a receptor determinant for infection of cells by avian Infectious bronchitis virus. *J. Gen. Virol.* 87, 1209–1216. <https://doi.org/10.1099/vir.0.81651-0>
- Woo, P.C.Y., Lau, S.K.P., Lam, C.S.F., Lai, K.K.Y., Huang, Y., Lee, P., Luk, G.S.M., Dyrting, K.C., Chan, K.-H., Yuen, K.-Y., 2009. Comparative analysis of complete genome sequences of three avian coronaviruses reveals a novel group 3c coronavirus. *J. Virol.* 83, 908–917. <https://doi.org/10.1128/JVI.01977-08>
- Woo, P.C.Y., Lau, S.K.P., Lam, C.S.F., Lau, C.C.Y., Tsang, A.K.L., Lau, J.H.N., Bai, R., Teng, J.L.L., Tsang, C.C.C., Wang, M., Zheng, B.-J., Chan, K.-H., Yuen, K.-Y., 2012. Discovery of Seven Novel Mammalian and Avian Coronaviruses in the Genus Deltacoronavirus Supports Bat Coronaviruses as the Gene Source of Alphacoronavirus and Betacoronavirus and Avian Coronaviruses as the Gene Source of Gammacoronavirus and Deltacoronavirus. *Journal of Virology* 86, 3995–4008. <https://doi.org/10.1128/JVI.06540-11>
- Woo, P.C.Y., Lau, S.K.P., Lam, C.S.F., Tsang, A.K.L., Hui, S.-W., Fan, R.Y.Y., Martelli, P., Yuen, K.-Y., 2014. Discovery of a novel bottlenose dolphin coronavirus reveals a distinct species of marine mammal coronavirus in Gammacoronavirus. *J. Virol.* 88, 1318–1331. <https://doi.org/10.1128/JVI.02351-13>
- Wu, C.-H., Chen, P.-J., Yeh, S.-H., 2014. Nucleocapsid Phosphorylation and RNA Helicase DDX1 Recruitment Enables Coronavirus Transition from Discontinuous to Continuous Transcription. *Cell Host Microbe* 16, 462–472. <https://doi.org/10.1016/j.chom.2014.09.009>
- Yamada, Y., Liu, D.X., 2009. Proteolytic activation of the spike protein at a novel RRRR/S motif is implicated in furin-dependent entry, syncytium formation, and infectivity of coronavirus infectious bronchitis virus in cultured cells. *J. Virol.* 83, 8744–8758. <https://doi.org/10.1128/JVI.00613-09>
- Yang, D., Leibowitz, J.L., 2015. The structure and functions of coronavirus genomic 3' and 5' ends. *Virus Res* 206, 120–133. <https://doi.org/10.1016/j.virusres.2015.02.025>
- Yang, X., Charlebois, P., Macalalad, A., Henn, M.R., Zody, M.C., 2013. V-Phaser 2: variant inference for viral populations. *BMC Genomics* 14, 674. <https://doi.org/10.1186/1471-2164-14-674>
- YuDong, W., YongLin, W., ZiChun, Z., GenChe, F., YiHai, J., XiangE, L., Jiang, D., ShuShuang, W., 1998. Isolation and identification of glandular stomach type IBV (QX IBV) in chickens. *Chinese Journal of Animal Quarantine* 15, 1–3.
- Zeng, Z., Deng, F., Shi, K., Ye, G., Wang, G., Fang, L., Xiao, S., Fu, Z., Peng, G., 2018. Dimerization of Coronavirus nsp9 with Diverse Modes Enhances Its Nucleic Acid Binding Affinity. *J Virol* 92. <https://doi.org/10.1128/JVI.00692-18>

Titre : *Gammacoronavirus* aviaires : dynamique évolutive des populations virales au cours d'infections expérimentales chez les poulets et les dindes.

Mots clés : *Gammacoronavirus* aviaires, Evolution, Séquençage Haut Débit

Résumé : Les deux principaux *Gammacoronavirus* aviaires, sont le virus de la bronchite infectieuse (IBV) du poulet et le coronavirus de la dinde (TCoV), deux virus génétiquement très proches, à l'exception de leur gène S. Malgré leur ressemblance, la diversité génétique au sein des deux virus est très différente. L'IBV présente une grande diversité génétique qui a été supposée comme étant le résultat d'une vaccination intensive des élevages de poulets qui aurait forcé l'évolution de l'IBV. En contraste, très peu de souches de TCoV ont été détectées et aucun vaccin n'est disponible. L'objectif de ce projet de thèse était donc d'analyser au cours de plusieurs passages ces deux virus dans leurs hôtes en analysant les variants génétiques présents au sein des différentes populations virales avec ou sans pression de sélection pour l'IBV et ce par séquençage haut débit. Pour cela, le développement d'un outil permettant la détection de variants génétiques viraux a été réalisé.

Par ailleurs, des méthodes ont été développées afin d'enrichir les échantillons en ARN viral dans le but de mieux les séquencer.

L'analyse des populations virales au cours des différents passages a permis de constater des changements rapides de séquence consensus. Pour IBV, une évolution différente du virus a été observée en fonction du statut immunitaire des sujets, avec une sélection de sept variants chez les sujets non vaccinés, contre un chez les sujets vaccinés. Pour TCoV, trois séries de dix passages ont été réalisées au cours desquels cinq changements de séquence consensus ont été observés dès le premier passage et se sont maintenus au cours des passages 5 et 10. Ces changements ont été retrouvés dans les différentes répliques expérimentales, indiquant donc que ces sélections ne sont pas aléatoires. En conclusion ces travaux montrent que l'évolution des *Gammacoronavirus* aviaires est rapide, dès le premier passage, chez leur hôte naturel.

Title : Avian *Gammacoronaviruses*: Evolutionary dynamics of viral populations during experimental infections in chickens and turkeys.

Keywords : Avian *Gammacoronaviruses*, Evolution, Next Generation Sequencing

Abstract : The infectious bronchitis virus (IBV) and the turkey coronavirus (TCoV), are the two major avian *Gammacoronaviruses*. Both are genetically related except for their S gene. Despite their similarity, the genetic diversity within the two viruses is different. IBV exhibits a huge genetic diversity which has been suggested to be the result of a massive vaccination of chicken flocks that would have forced the evolution of IBV. In contrast, very few TCoV strains have been detected and no vaccination of turkeys was performed because no vaccine is available. The objective of this thesis project was to analyze the evolution of these two avian coronaviruses in their hosts during several passages by analysing the genetic variants present in the viral population using NGS with or without selection pressure for IBV.

To this purpose, a bioinformatic pipeline that allows the detection of variants has been developed.

Also, methods have been developed to improve viral RNA sequencing. The analysis of the viral populations during the different passages revealed rapid changes in consensus sequence of IBV and TCoV. For IBV, viral evolution was different depending on the immune status of the subjects, with a selection of seven variants in the unvaccinated subjects, against one in the vaccinated subjects. For TCoV, three series of ten passages were performed in which five modifications of consensus sequence were detected as soon as the first passage and were retrieved during passages 5 and 10. The changes were found in the different experimental replicates, indicating that these selections are not random. In conclusion, this work shows that the evolution of avian *Gammacoronaviruses* is rapid as soon as the first passage, in their natural host.